

CERTIFIED COPY OF
PRIORITY DOCUMENT

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

M. Nishihara
5/30/00
Q 59233
10fu

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

JC625 U.S. PTO
09/580559
05/30/00

出 願 年 月 日
Date of Application:

1999年 6月 1日

願 番 号
Application Number:

平成11年特許願第154308号

願 人
Applicant(s):

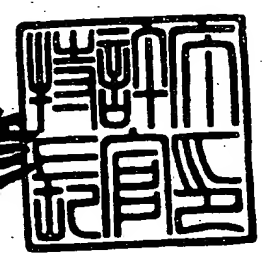
日本電気株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 3月17日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



【書類名】 特許願

【整理番号】 49210375

【提出日】 平成11年 6月 1日

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/56
H04Q 3/00

【発明の名称】 パケット転送方法および装置ならびにパケット通信システム

【請求項の数】 22

【発明者】

【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

【氏名】 西原 基夫

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100108578

【弁理士】

【氏名又は名称】 高橋 詔男

【代理人】

【識別番号】 100064908

【弁理士】

【氏名又は名称】 志賀 正武

【選任した代理人】

【識別番号】 100101465

【弁理士】

【氏名又は名称】 青山 正和

【選任した代理人】

【識別番号】 100108453

【弁理士】

【氏名又は名称】 村山 靖彦

【選任した代理人】

【識別番号】 100100077

【弁理士】

【氏名又は名称】 大場 充

【手数料の表示】

【予納台帳番号】 008707

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9709418

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット転送方法および装置ならびにパケット通信システム

【特許請求の範囲】

【請求項 1】 パケットを送受信するアクセス網と、前記パケットを中継する中継手段を備えたバックボーンとなるネットワークとの間で前記パケットを転送するパケット転送装置であって、

前記パケットを前記中継手段のスイッチ単位である固定長セルの n 倍 (n は 2 以上の整数) の長さを持つスーパーパケットに変換して前記ネットワークに送出する構築手段と、

前記ネットワークから送られてくる前記スーパーパケットより前記パケットを取り出して前記アクセス網に送出する復元手段と

を具備することを特徴とするパケット転送装置。

【請求項 2】 前記構築手段は、前記パケットの宛先となるアクセス網と前記ネットワークとの間に配置されたパケット転送装置毎に前記パケットを蓄積するキューが作られる第 1 の記憶手段を具備し、前記パケットを宛先毎に該キューに蓄積してゆき、前記スーパーパケットの構築に必要なだけのパケットが蓄積されたことを検出し、該キューから前記パケットを抽出して前記スーパーパケットを構築することを特徴とする請求項 1 記載のパケット転送装置。

【請求項 3】 前記構築手段は、前記スーパーパケットが構築される度に計時を開始して該計時開始時点から第 1 のタイムアウト時間が経過したことを検出する第 1 の計時手段を前記第 1 の記憶手段上のキュー毎に具備し、該第 1 のタイムアウト時間の経過が検出された時点で該キューに蓄積されているパケットから前記スーパーパケットを構築することを特徴とする請求項 2 記載のパケット転送装置。

【請求項 4】 前記第 1 のタイムアウト時間は、予め決められた前記ネットワーク上のトラヒック毎の最小帯域から決定されることを特徴とする請求項 3 記載のパケット転送装置。

【請求項 5】 前記第 1 のタイムアウト時間は、予め決められた前記ネットワークにおけるトラヒック毎の許容網内遅延時間から決定されることを特徴とす

る請求項 3 記載の packets 転送装置。

【請求項 6】 前記復元手段は、送信元のアクセス網と前記ネットワークとの間に配置され前記スーパー packets を送出する packets 転送装置毎に前記スーパー packets を蓄積するキューが作られる第 2 の記憶手段を具備し、前記スーパー packets を該キューに蓄積してゆき、蓄積された前記スーパー packets から前記 packets を復元することを特徴とする請求項 1 ～ 5 の何れかの項記載の packets 転送装置。

【請求項 7】 前記復元手段は、前記 packets が復元される度に計時を開始して該計時開始時点から第 2 のタイムアウト時間が経過したことを検出する第 2 の計時手段を前記第 2 の記憶手段上のキュー毎に具備し、該第 2 のタイムアウト時間の経過が検出された時点で該キューに滞留しているスーパー packets を廃棄することを特徴とする請求項 6 記載の packets 転送装置。

【請求項 8】 前記第 2 のタイムアウト時間は、予め決められた前記ネットワーク上のトラヒックの最小帯域または網内遅延の最大許容値、前記ネットワーク内の遅延分散時間、及び、予め定められている保護時間から決定されることを特徴とする請求項 7 記載の packets 転送装置。

【請求項 9】 前記構築手段は、前記 packets が複数のスーパー packets にまたがることを検出してこれら複数のスーパー packets に分割して格納し、

前記復元手段は、前記スーパー packets 上の packets が複数のスーパー packets にまたがっていることを検出し、これら複数のスーパー packets に分割して格納されている packets のデータを繋げて送信元のアクセス網が生成した元の packets を復元することを特徴とする請求項 1 ～ 8 の何れかの項記載の packets 転送装置。

【請求項 10】 前記スーパー packets には、そのペイロードが単一の packets のデータで占有されているか否かを示す単一 packets 占有情報が格納され、

前記構築手段は、前記 packets の各々の長さを前記スーパー packets のペイロードの長さと比較し、該比較結果に応じて前記単一 packets 占有情報を占有有又は占有無に設定し、

前記復元手段は、前記スーパー packets 上の前記単一 packets 占有情報が占有

有であることを検出し、前記単一パケット占有情報が占有無のスーパーパケットが到来するまでのこれら一連のスーパーパケット上に載せられているパケットのデータを繋げて送信元のアクセス網が生成した元のパケットを復元することを特徴とする請求項 1～9 の何れかの項記載のパケット転送装置。

【請求項 11】 前記スーパーパケットのヘッダは前記パケットのヘッダと同一のフォーマットを有していることを特徴とする請求項 1～10 の何れかの項記載のパケット転送装置。

【請求項 12】 パケットを送受信するアクセス網と、

請求項 1～11 の何れかの項記載のパケット転送装置と、

パケットを中継する中継手段を備え、該中継手段のスイッチ単位である固定長セルの n 倍 (n は 2 以上の整数) の長さを持つスーパーパケットを転送するバックボーンのネットワークとを具備し、

前記パケット転送装置は、前記アクセス網によって送受信される前記パケットと前記ネットワーク上で転送される前記スーパーパケットとを相互に変換して、送信元のアクセス網から送信される前記パケットを前記スーパーパケットの形で前記ネットワーク内の前記中継手段を経由させて宛先のアクセス網まで転送することを特徴とするパケット通信システム。

【請求項 13】 前記パケット転送装置及び前記中継手段には、前記ネットワーク内でのみ定義される固有の網内アドレスがそれぞれ付与されており、

前記スーパーパケットのヘッダには、送信元アドレス及び宛先アドレスとして、送信元及び宛先のアクセス網にそれぞれ接続されたパケット転送装置の網内アドレスが格納されることを特徴とする請求項 12 記載のパケット通信システム。

【請求項 14】 宛先となるパケット転送装置が同一であるスーパーパケットのうち、該宛先のパケット転送装置に接続された宛先のアクセス網の種別に応じて、該スーパーパケットのヘッダ中の宛先アドレスに異なる網内アドレスが割り当てられていることを特徴とする請求項 13 記載のパケット通信システム。

【請求項 15】 前記中継手段は、

前記スーパーパケットのヘッダ中の宛先アドレスと該スーパーパケットの中継先を対応づけたエントリが、前記パケット転送装置および前記中継手段の数に相

当するだけ少なくとも格納された経路検索テーブルと、

前記スーパーパケットのヘッダ中の宛先アドレスをもとに前記経路検索テーブルを検索して該スーパーパケットの中継先を特定する経路検索手段とを具備することを特徴とする請求項 13 又は 14 記載のパケット通信システム。

【請求項 16】 送信元のアクセス網に接続されたパケット転送装置から前記ネットワーク内の各中継手段を経て宛先のアクセス網に接続されたパケット転送装置に至るまでのトラヒックである網内フローが同一であれば、該同一の網内フローに対応するスーパーパケットのヘッダを互いに同じにしていることを特徴とする請求項 12～15 の何れかの項記載のパケット通信システム。

【請求項 17】 送信元のアクセス網が、宛先のアクセス網に転送すべきパケットを生成して自身に接続された送信元パケット転送装置へ送信し、

該送信元パケット転送装置は、前記送信元のアクセス網から送信されたパケットをバックボーンとなるネットワーク上に設けられた中継手段のスイッチ単位である固定長セルの n 倍 (n は 2 以上の整数) の長さを持つスーパーパケットに変換して前記ネットワークに送出し、

前記ネットワークは、前記中継手段を用いて前記スーパーパケットを中継して前記宛先のアクセス網に接続された宛先パケット転送装置に至るまで転送し、

該宛先パケット転送装置は、前記ネットワークから転送される前記スーパーパケットをもとに前記送信元のアクセス網により生成された前記パケットを復元して前記宛先のアクセス網に送出することを特徴とするパケット転送方法。

【請求項 18】 前記送信元パケット転送装置は、前記送信されたパケットを前記宛先パケット転送装置を単位として個別に蓄積し、該宛先パケット転送装置毎に前記スーパーパケットを構築して前記ネットワークに送出し、

前記宛先パケット転送装置は、前記ネットワークから転送されてくる前記スーパーパケットを前記送信元パケット転送装置を単位として個別に蓄積し、該送信元パケット転送装置毎に前記パケットの復元を行うことを特徴とする請求項 17 記載のパケット転送方法。

【請求項 19】 前記送信元パケット転送装置は、前記スーパーパケットが第 1 のタイムアウト時間にわたって構築されていないことを前記宛先パケット転

送装置毎に検出し、該構築が行われていない宛先パケット転送装置に関して蓄積されているパケットが存在していれば、該パケットから前記スーパーパケットを構築して前記ネットワークに送出することを特徴とする請求項 1 8 記載のパケット転送方法。

【請求項 2 0】 前記宛先パケット転送装置は、前記パケットが第 2 のタイムアウト時間にわたって復元されていないことを前記送信元パケット転送装置毎に検出し、該復元が行われていない送信元パケット転送装置に関して蓄積されているスーパーパケットが存在していれば、該スーパーパケットを廃棄することを特徴とする請求項 1 8 又は 1 9 記載のパケット転送方法。

【請求項 2 1】 前記送信元パケット転送装置は、前記送信元のアクセス網から送信された前記パケットが複数のスーパーパケットにまたがる場合に該パケットをこれら複数のスーパーパケットに分割して前記ネットワークに送出し、

前記宛先パケット転送装置は、前記ネットワークから転送されてくる前記スーパーパケット中のパケットが複数のスーパーパケットにまたがっている場合に、これら複数のスーパーパケットに分割されているパケットのデータを繋げて前記送信元のアクセス網によって生成された元のパケットを復元して前記宛先のアクセス網に送出することを特徴とする請求項 1 7 ～ 2 0 の何れかの項記載のパケット転送方法。

【請求項 2 2】 前記送信元パケット転送装置は、前記スーパーパケットのヘッダ中の送信元アドレス及び宛先アドレスとして、該送信元パケット転送装置及び前記宛先パケット転送装置にそれぞれ付与されている前記ネットワーク内でのみ定義される固有の網内アドレスを格納して前記ネットワークに送出し、

前記ネットワーク内の各中継手段は、転送されてくる前記スーパーパケットのヘッダを参照して前記宛先パケット転送装置に付与された網内アドレスに従って該スーパーパケットの中継先を特定して該スーパーパケットを前記宛先パケット転送装置まで転送することを特徴とする請求項 1 7 ～ 2 1 の何れかの項記載のパケット転送方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、ネットワーク上でパケットやフレームを転送するためのパケット転送方法および装置ならびにこのパケット転送装置を利用したパケット通信システムに関し、特に、ネットワーク上で転送可能なパケットの数を増大させるための技術に関するものである。

【0002】

【従来の技術】

(1) 従来のネットワークの構成

図15は従来から用いられているパケット通信のためのネットワークの構成を示したブロック図である。ここでは、パケット転送装置としてIP（インターネット・プロトコル）パケットを転送対象としたルータ装置を例に挙げて説明する。図示したように、ネットワーク内にはパケットを中継するためのルータ装置が複数台配置されており、同図では4台のルータ装置1a～1dを例示してある。これらルータ装置の間は互いに伝送路を介して接続されており、例えばルータ装置1aとルータ装置1bとの間は伝送路2abで接続され、これら以外のルータ装置間でも同様にして伝送路2ac, 2ad, 2bc, 2bd, 2cdを介してそれぞれ接続されている。

【0003】

各ルータ装置にはアクセス網の一例である1ないし複数のユーザLAN（ローカル・エリア・ネットワーク）が接続されており、ルータ装置とこれらユーザLANとの間がアクセス回線によって接続されている。同図では各ルータ装置に対してそれぞれ2つのユーザLANを接続した形態について例示してあり、例えばルータ装置1aにはユーザLAN3a1, 3a2がそれぞれアクセス回線4a1, 4a2を介して接続されている。このルータ装置1a以外の各ルータ装置についても同様であってユーザLAN3b1, 3b2, 3c1, 3c2, 3d1, 3d2がそれぞれアクセス回線4b1, 4b2, 4c1, 4c2, 4d1, 4d2を介して接続されている。また、各ユーザLANには1台ないし複数台のホスト端末が接続されており、同図ではユーザLAN3a1, 3d1にそれぞれ接続された送信元ホスト5a1, 宛先ホスト5d1のみを示すこととし、その他のホス

ト端末は図示を省略してある。

【0004】

いま、図15の構成において、送信元ホスト5a1から宛先ホスト5d1に対して転送ルート6を通じてIPパケットを送信する場合を想定する。なお、伝送路上およびアクセス回線上ではIPパケット等がいずれもレイヤ1フレームの形で転送されるが、煩雑になることから、以下の説明中では「IPパケットが伝送路やアクセス回線上を転送される」といった表現を用いることがある。そして、転送ルート6を利用した場合、送信元ホスト5a1が発生させたIPパケットは、レイヤ1フレームの形にされてユーザLAN3a1およびアクセス回線4a1を通じてルータ装置1aに送られる。

【0005】

ルータ装置1aは、送られたレイヤ1フレーム中のIPパケットのヘッダに記述されている宛先アドレスを元に当該IPパケットの転送経路を判定して転送すべきルータ装置又はユーザLANを特定し、これと同時に、特定されたルータ装置又はユーザLANに対応した転送経路としての伝送路又はアクセス回線を特定する。同図の例では、当該IPパケットの転送先が伝送路2ad、ルータ装置1dに特定されるので、ルータ装置1aは伝送路2adに対してIPパケットの含まれたレイヤ1フレームを送出する。これにより、ルータ装置1dはルータ装置1aと同様にIPパケットを転送してゆき、その結果、上記IPパケットはアクセス回線4d1を介して宛先のユーザLAN3d1に転送されて、最終的にこのユーザLAN内の宛先ホスト5d1にまで転送される。

【0006】

(2) 従来のルータ装置の構成

次に、図15に示した各ルータ装置の構成および動作に関して説明する。図16はルータ装置の構成を示したブロック図であって、1ないし複数の入力インタフェース部10-1～10-N（N：自然数），1ないし複数の出力インタフェース部11-1～11-N，スイッチスケジューラ部12，クロスバスイッチ部13から構成されている。このうち、各々の入力インタフェース部はレイヤ1/2処理部14，IPレイヤ処理部15，パケット分割部16，経路検索テーブル17よ

り構成されている。また、入力インタフェース部 10-1~10-Nには固有の入力インタフェース番号としてそれぞれ“1”~“N”が付与されている。一方、各々の出力インタフェース部はパケット復元部 18, IPレイヤ処理部 19, レイヤ 1/2 処理部 20 より構成されている。また、出力インタフェース部 11-1~11-Nには固有の出力インタフェース番号としてそれぞれ“1”~“N”が付与されている。

【0007】

各入力インタフェース部において、レイヤ 1/2 処理部 14 には図 15 に示した伝送路またはアクセス回線の何れかが接続される。例えば、入力インタフェース部 10-1 のレイヤ 1/2 処理部 14 が図 15 に示したアクセス回線 4a2 に接続され、入力インタフェース部 10-2 のレイヤ 1/2 処理部 14 が伝送路 2ac に接続されるなどである。このレイヤ 1/2 処理部 14 は、自身の属する入力インタフェース部で終端すべき伝送路やアクセス回線のレイヤ 1/2 プロトコルを終端する機能を有しており、終端された IP パケットを IP レイヤ処理部 15 に転送する。

【0008】

IP レイヤ処理部 15 は、レイヤ 1/2 処理部 14 から転送されてくる IP パケットのヘッダを参照して IP レイヤの処理を行う。なお、この IP レイヤ処理の詳細については本発明の要旨と関係ないことと、既存の技術であるためここではその説明を割愛する。また、パケット分割部 16 は、IP レイヤ処理部 15 から転送されてくる IP パケットをルータ装置内で定義される固定長の装置内セルに分割し、これら装置内セルをクロスバースイッチ部 13 に送出する。つまり、この固定長がルータ装置のスイッチング単位の長さに相当することになる。また、これと同時にパケット分割部 16 は、転送された IP パケットを出力インタフェース部 11-1~11-N の何れかに転送するために、スイッチスケジューラ部 12 に対して転送リクエスト信号 21 を送出する。

【0009】

スイッチスケジューラ部 12 はクロスバースイッチ部 13 に対するスイッチ設定を行う機能ブロックであって、双方向の信号線を介して、各入力インタフェー

ス部 10-1~10-N 内のパケット分割部 16 との間で転送リクエスト信号 21 (前述) および転送リプライ信号 22 を授受する。すなわち、パケット分割部 16 からスイッチスケジューラ部 12 に送出されるのが転送リクエスト信号 21 であって、これとは逆に、スイッチスケジューラ部 12 からパケット分割部 16 に送出されるのが転送リプライ信号 22 である。そしてスイッチスケジューラ部 12 は、各パケット分割部 16 から送信される転送リクエスト信号 21 を受け取り、クロスバースイッチ部 13 における最適なスイッチ設定を計算してこのクロスバースイッチ部 13 へスイッチ設定信号 23 を送る。

【0010】

クロスバースイッチ部 13 は、送出されたこのスイッチ設定信号 23 に基づいて、入力インタフェース部 10-1~10-N の何れかと出力インタフェース部 11-1~11-N の何れかとの間を接続する。各出力インタフェース部 11-1~11-N において、パケット復元部 18 はクロスバースイッチ部 13 を通じて各入力インタフェース部 10-1~10-N の何れかから転送されてくる固定長の装置内セルを積み上げて、元の IP パケット (即ち、レイヤ 1/2 処理部 14 で終端された IP パケット) を復元して IP レイヤ処理部 19 に転送する。IP レイヤ処理部 19 は、パケット復元部 18 から転送されてくる IP パケットのヘッダを参照して IP レイヤに関わる処理を行ってからレイヤ 1/2 処理部 20 へ転送する。

【0011】

レイヤ 1/2 処理部 20 は、転送されたこの IP パケットに対してレイヤ 1/2 のヘッダをそれぞれ付与してレイヤ 1 フレームを構築し、伝送路またはアクセス回線である出力方路に転送する。つまり、各出力インタフェース部のレイヤ 1/2 処理部 20 には、入力インタフェース部 10-1~10-N と同様に図 15 に示した伝送路またはアクセス回線の何れかが接続される。例えば、出力インタフェース部 11-1 のレイヤ 1/2 処理部 20 が図 15 に示したアクセス回線 4a2 に接続され、出力インタフェース部 11-2 のレイヤ 1/2 処理部 20 が伝送路 2ac に接続されるなどである。

【0012】

(3) 従来のルータ装置の動作

次に、図 16 に示したルータ装置の詳細動作について具体的に説明する。ここで、図 15 に示した伝送路やアクセス回線上を伝送される信号フォーマットは図 17 に示したような形式をしている。すなわち、前述したように伝送路アクセス回線上ではレイヤ 1 フレーム 30 が伝送される。このレイヤ 1 フレーム 30 にはレイヤ 1 ヘッダ 31 が周期的に挿入され、このレイヤ 1 ヘッダ 31 によってレイヤ 1 の機能である bit 同期, byte 同期, フレーム同期, 回線の状態監視やエラー率の監視などに使用される OAM (Operation Administration and Maintenance) 機能がレイヤ 1 / 2 処理部 14 によって実現される。

【0013】

また、レイヤ 1 ペイロード 32 にはレイヤ 2 フレーム 33 が詰め込まれており、この部分を介して可変長のレイヤ 2 フレーム 33 が転送される。レイヤ 2 フレーム 33 はレイヤ 2 ヘッダ 34 およびレイヤ 2 トレーラ 35 でレイヤ 2 ペイロード 36 を挟むようにして構成されており、レイヤ 2 ペイロード 36 の部分には IP パケット 37 が乗せられる。この IP パケット 37 はヘッダ 38 およびペイロード 39 から構成される。

【0014】

図 16 に示したレイヤ 1 / 2 処理部 14 は、伝送路やアクセス回線から入力される信号に対してレイヤ 1 処理を行い、当該処理によって得られたレイヤ 2 フレーム 33 に含まれるレイヤ 2 ヘッダ 34 およびレイヤ 2 トレーラ 35 を参照してレイヤ 2 処理を実行する。なお、レイヤ 1 処理およびレイヤ 2 処理はいずれも既存の技術であって、本発明の要旨とは直接関係しないのでここでは説明を割愛する。その後、レイヤ 1 / 2 処理部 14 はレイヤ 2 ペイロード 36 に含まれた IP パケット 37 を取り出してこれを IP レイヤ処理部 15 に転送する。

【0015】

IP レイヤ処理部 15 は転送された IP パケット 37 のヘッダ 38 を参照して、IP ヘッダの正当性チェックや異常な IP パケット受信処理等の IP レイヤ処理を実行するほか、ヘッダ 38 に含まれている宛先アドレスフィールドの内容に従って経路検索テーブル 17 内の経路情報を検索する。この検索動作によって、IP レイヤ処理部 15 は次に IP パケットを転送すべきルータ装置またはユーザ

LANを特定するとともに、これらルータ装置やユーザLANに対応した出力インタフェース部11-1～11-Nの何れかに付与されている出力インタフェース番号（すなわち、“1”～“N”）を特定し、パケット分割部16に対してIPパケット（図16では「データ信号」としてある）および出力インタフェース番号を転送する。

【0016】

パケット分割部16は、IPレイヤ処理部15から転送されてくるIPパケット37を図17に示したように固定長の装置内セル40に分割してペイロード部分に収容するとともに、セルヘッダ41を装置内セル40毎に付与する。ここで、セルヘッダ41には、転送先の出力インタフェース番号のほか、送信元である自身の入力インタフェース部に付与されている入力インタフェース番号（例えば、入力インタフェース部10-2内のパケット分割部16であれば“2”）が書き込まれる。

【0017】

また、パケット分割部16は出力インタフェース番号単位に異なるメモリ（図示省略）を内蔵しており、分割された装置内セル40を出力インタフェース番号に対応するメモリへ格納する。さらに、パケット分割部16は上記メモリ内に格納されている装置内セル40をこれらのセルヘッダ41に書き込まれた出力インタフェース番号に相当する出力インタフェース部11-1～11-Nの何れかへ転送すべく、転送リクエスト信号21をスイッチスケジューラ部12に送出する。

【0018】

スイッチスケジューラ部12は、各入力インタフェース部10-1～10-Nから送出される転送リクエスト信号21に対して、入力インタフェース部－出力インタフェース部間におけるスイッチ接続の適性な組み合わせを計算する。次に、スイッチスケジューラ部12は、IPパケット37を出力インタフェース部11-1～11-Nの何れに転送すべきかを各入力インタフェース部10-1～10-Nへ通知するために、これら入力インタフェース部に対して転送リプライ信号22を返送する。この転送リプライ信号22を受け取ると、各入力インタフェース部10-1～10-Nは当該信号によって指定された出力インタフェース部に向けて固定長の

装置内セル40をクロスバースイッチ部23に順次転送してゆく。これと同時にスイッチスケジューラ部12は、クロスバースイッチ部13に対してスイッチ設定信号23を送出する。

【0019】

クロスバースイッチ部13は送られたスイッチ設定信号23を元にして、指定された入力インタフェース部と指定された出力インタフェース部との間でスイッチ接続を行う。以上のようにして、入力インタフェース部10-1～10-Nから出力インタフェース部11-1～11-Nに対して装置内セル40が順に転送されてゆく。すると、指定された出力インタフェース部内のパケット復元部18は、指定された入力インタフェース部から転送されてくる装置内セル40をセルヘッダ41に記述されている入力インタフェース番号に基づいて異なる内蔵メモリ（図示省略）へ格納する。また、パケット復元部18は装置内セル40を積み上げてゆき、これら装置内セル40からセルヘッダ41を除いたペイロード部分から元のIPパケット37を復元してIPレイヤ処理部19に転送する。

【0020】

IPレイヤ処理部19は転送されたIPパケット37に対してIPレイヤ処理を行った後にレイヤ1/2処理部20に転送する。レイヤ1/2処理部20は転送されたIPパケット37をレイヤ2ペイロード36に収容し、その前後にそれぞれレイヤ2ヘッダ34およびレイヤ2トレーラ35を付与してレイヤ2フレーム33を構築する。次に、レイヤ1/2処理部20は構築されたレイヤ2フレーム33をレイヤ1ペイロード32に収容してその前にレイヤ1ヘッダ31を付与してレイヤ1フレーム42を構築したのち、当該レイヤ1フレーム42を指定された出力インタフェース部に接続されている伝送路またはアクセス回線に送出する。

【0021】

以上の動作によって、伝送路またはアクセス回線からこれらに対応した入力インタフェースを介してルータ装置に転送されてくるIPパケットが、該当する出力インタフェースに対応した伝送路またはアクセス回線に転送されてゆく。

なお、以上のような技術に類似するものとして次のようなものがある。例えば

、上記の説明では可変長の packets を固定長のセルに分割していたのに対し、短 packets を固定長の packets (53 バイトの ATM [非同期転送モード] セル) に多重化してネットワーク上のノード間を転送させるものとして、再公表特許 WO/26589 号公報, 特開平 7-245628 号公報などがある。また、宛先が同一である短 packets を長 packets にまとめてネットワーク上を転送させるものとして、特開昭 63-197148 号公報などがある。

【0022】

【発明が解決しようとする課題】

次に、上述した従来のルータ装置を用いた IP packets 転送に存在する問題点について説明する。

(1) 従来の第 1 の問題点

前述したように、図 16 に示したクロスバースイッチ部 13 は図 17 に示したような固定長の装置内セル 40 を対象としたスイッチである。このため、装置内セル 40 の転送時間単位を「 τ 」と置いたときに、クロスバースイッチ部 13 のスイッチ接続動作をこの転送時間単位 τ で切り替えてやる必要がある。したがって、スイッチスケジューラ部 12 はクロスバースイッチ部 13 で行われるスイッチ接続の設定を転送時間単位 τ 内で毎回計算しなければならないという制約が生じることになる。このことは、上述した再公表特許 WO/26589 号公報, 特開平 7-245628 号公報などでも同様であって、これら公報では 53 バイトと短い ATM セルを packets スイッチでスイッチングさせている。

【0023】

ここで、ルータ装置の持つスイッチ容量を増加させるためには、入力インタフェース部 10-1~10-N や出力インタフェース部 11-1~11-N について、インタフェース速度を変えずにインタフェース数を増やすか、あるいは、インタフェース数を変えずにインタフェース速度を増やすかのいずれかの対処が必要となる。いま前者で対処するとした場合、入力インタフェース部や出力インタフェース部の数が増加するため、それにつれてスイッチ接続の設定のための計算量が増えてゆくことになる。しかるに、ルータ装置を実現しているデバイスの能力にも当然ながら限界があることから、転送時間単位 τ の範囲内で処理できる計算量 (換

言すれば、スイッチスケジューラ部 1 2 の計算能力) にも上限がある。このため、ある上限を越えてインタフェース数を増やすことはできず、スイッチ容量を一定以上に増やすこともできないという問題がある。

【0 0 2 4】

一方、上述した 2 つの選択肢のうちの後者で対処するとした場合、入力インターフェイス部や出力インタフェース部のインタフェース速度を上げると装置内セル 4 0 の転送時間単位 τ がそれに反比例して短くなってしまふ。これはスイッチ接続設定のための計算時間が短くなることと等価であつて、インタフェース速度を上げる前と比べると、規定された転送時間単位 τ 内ではより少ないインタフェース数についてしかスイッチ接続設定の計算を実行できなくなる。つまり、この場合もやはりスイッチ容量を大きくすることが困難になるという問題が出てくる。以上のように、装置内セル 4 0 の長さによってスイッチ容量ないしはスイッチ規模の上限が決定されるため、ルータ装置のスイッチ容量を一定以上に増やせないという問題がある。

【0 0 2 5】

(2) 従来の第 2 の問題点

第 1 の問題点のところでも述べたように、ルータ装置の内部でスイッチされる固定長の装置内セル 4 0 が小さくセル転送時間が短いと、スイッチスケジューラ部 1 2 で行われるスイッチ接続設定のための計算時間の限界を画する転送時間単位 τ が短くなる (これは、上述した再公表特許 W O / 2 6 5 8 9 号公報, 特開平 7 - 2 4 5 6 2 8 号公報などでも同様)。このため、どうしても装置内セル 4 0 の長さを或る一定長以上に大きくせざるを得ない。特に、大規模なルータ装置では、I P パケットの最小長である 6 4 b y t e 以下の装置内セルを定義してしまうとスイッチ接続の設定に必要な計算時間を確保できなくなつてしまふ。このため、現存する大規模ルータでは装置セルの長さを 1 2 8 b y t e や 2 5 6 b y t e 等として、I P パケットの最小長よりも長い装置内セルを定義していることが多い。そのため、転送されてくる I P パケットの長さの分布によってはスイッチの際のスループットが著しく落ちてしまふなどの問題が生じてくる。

【0 0 2 6】

この点について以下にさらに詳しく説明する。いま仮に装置内セル40の大きさを128byteとし、各入力インタフェース部10-1~10-Nに対して最小長（即ち、64byte長）のIPパケットが連続して入力され、なおかつ、これらIPパケットの各々が出力インタフェース部11-1~11-Nのうちの異なる出力インタフェース部に転送されるという状況を想定する。ここで、図16に示したパケット分割部16およびパケット復元部18が設けられた本来の目的は、可変長のIPパケットをルータ装置内で使用される装置内セル40に分割し、分割されたこれら装置内セル40をクロスバスイッチ部13でスイッチさせ、出力インタフェース部11-1~11-Nにてスイッチされた装置内セル40から再びIPパケットを組み立て直すことにあったはずである。

【0027】

しかしながら、上述したような状況が想定される場合にはIPパケット自体が装置内セルより短いため、IPパケットを分割することができず、IPパケットそのものを単位として装置内セル40を構築して転送せざるを得ない。このため、装置内セル40のペイロード領域の一部が未使用となるが、この未使用領域がそのままスイッチの未使用帯域になってしまうため、ルータ装置におけるスイッチ動作のスループットが上記未使用帯域の量だけ低下してしまうことになる。例えば図18に示したように、64byte長のパケットPA、PB、PCがルータ装置の入力インタフェース部の何れかに順次入力されて、入力インターフェース部内のパケット分割部16に伝送されたものとする。

【0028】

このとき、これらIPパケットはそれぞれ異なる出力インタフェース部に転送されることと、装置内セル40の大きさは図にも示したように128byteであるため、IPパケットを装置内セル40に挿入してもその半分の64byte（図中の「空き帯域」に相当）は未使用になってしまう。また、パケット分割部16に対してIPパケットが入力される入力速度と、パケット分割部16から装置内セル40が出力される出力速度は同じであるため、入力された情報量の半分の情報量に相当するだけの装置内セルしか出力できないことになる。このため、パケット分割部16に入力されたIPパケットのうちの半分はいずれパケット分

割部 16 で廃棄されることになるため、スイッチ動作のスループットが 50% に減少してしまうという問題がある。なお、上述した特開昭 63-197148 号公報においても、パケット交換網内のルーティング処理手段が実際には長パケットを装置内セルに分解してスイッチングさせていると考えられるため、上記と同様の問題が発生することになる。

【0029】

(3) 従来の第 3 の問題点

一般に IP パケットは可変長であるため、ルータ装置内のパケット分割部 16 で IP パケットを固定長の装置内セル 40 に分割すると、IP パケットの最後尾が含まれた装置内セルに空きが発生することになる。例えば、図 19 に示したように単一の IP パケット 37 を装置内セル 40a~40f に分割してゆくと、最後尾にある装置内セル 40f に空き（図中の「空き領域」に相当）が生じる。このため、この空きの部分については Padding を行う必要が生じてくる。したがって、これら一連の装置内セル 40a~40f をクロスバースイッチ部 13 に転送すると、最後の装置内セル 40f に含まれる空きが必ず未使用帯域となってスイッチ転送容量の減少を引き起こしてしまう。

【0030】

例えばいま装置内セルの長さが 128 byte であるという条件で 129 byte の IP パケットばかり発生する状況を想定すると、スイッチ転送容量は $129 / (128 \times 2) \approx 50\%$ にまで低下してしまうことになる。なお、このような問題は、ATM セル単位でスイッチングを行っている上述の再公表特許 WO/26589 号公報および特開平 7-245628 号公報などや、パケット交換網内のルーティング処理手段が実際には長パケットを装置内セルに分解してスイッチングさせていると考えられる上述の特開昭 63-197148 号公報でも同様に発生する。

【0031】

本発明は上記の点に鑑みてなされたものであり、その目的は、パケット転送装置を構成しているクロスバースイッチ部のスイッチ単位ないし転送時間単位の大きさを増やすことができるほか、入出力インタフェース数や入出力インタフェー

ス速度が増大してもスイッチ接続設定に要する計算時間を確保することができ、スイッチ容量を大幅に増やしてネットワーク上で転送可能なパケット数を増やせる大規模なパケット転送装置と、このパケット転送装置を利用して構築されるパケット通信システムを提供することにある。

【0032】

また、本発明の目的は、ネットワーク上で転送されるパケットの最小長を十分大きくすることができ、空き帯域の発生によるスイッチ動作のスループット低下を来すことなく、ネットワーク上で転送可能なパケット数を増やすことのできるパケット転送装置およびパケット通信システムを提供することにある。

また、本発明の目的は、ネットワーク上で転送されるパケットの長さを装置内セルの長さの整数倍として、パケットを装置内セル単位に分割してスイッチ動作させた場合にも無駄な空き帯域が生じることなく、ネットワーク上で転送可能なパケット数を増やせるパケット転送装置およびパケット通信システムを提供することにある。

【0033】

【課題を解決するための手段】

以上の課題を解決するため、請求項1記載の発明は、パケットを送受信するアクセス網と、前記パケットを中継する中継手段を備えたバックボーンとなるネットワークとの間で前記パケットを転送するパケット転送装置であって、前記パケットを前記中継手段のスイッチ単位である固定長セルの n 倍（ n は2以上の整数）の長さを持つスーパーパケットに変換して前記ネットワークに送出する構築手段と、前記ネットワークから送られてくる前記スーパーパケットより前記パケットを取り出して前記アクセス網に送出する復元手段とを具備することを特徴としている。

また、請求項2記載の発明は、請求項1記載の発明において、前記構築手段は、前記パケットの宛先となるアクセス網と前記ネットワークとの間に配置されたパケット転送装置毎に前記パケットを蓄積するキューが作られる第1の記憶手段を具備し、前記パケットを宛先毎に該キューに蓄積してゆき、前記スーパーパケットの構築に必要なだけのパケットが蓄積されたことを検出し、該キューから前

記パケットを抽出して前記スーパーパケットを構築することを特徴としている。

また、請求項3記載の発明は、請求項2記載の発明において、前記構築手段は、前記スーパーパケットが構築される度に計時を開始して該計時開始時点から第1のタイムアウト時間が経過したことを検出する第1の計時手段を前記第1の記憶手段上のキュー毎に具備し、該第1のタイムアウト時間の経過が検出された時点で該キューに蓄積されているパケットから前記スーパーパケットを構築することを特徴としている。

【0034】

また、請求項4記載の発明は、請求項3記載の発明において、前記第1のタイムアウト時間は、予め決められた前記ネットワーク上のトラヒック毎の最小帯域から決定されることを特徴としている。

また、請求項5記載の発明は、請求項3記載の発明において、前記第1のタイムアウト時間は、予め決められた前記ネットワークにおけるトラヒック毎の許容網内遅延時間から決定されることを特徴としている。

また、請求項6記載の発明は、請求項1～5の何れかの項記載の発明において、前記復元手段は、送信元のアクセス網と前記ネットワークとの間に配置され前記スーパーパケットを送出するパケット転送装置毎に前記スーパーパケットを蓄積するキューが作られる第2の記憶手段を具備し、前記スーパーパケットを該キューに蓄積してゆき、蓄積された前記スーパーパケットから前記パケットを復元することを特徴としている。

また、請求項7記載の発明は、請求項6記載の発明において、前記復元手段は、前記パケットが復元される度に計時を開始して該計時開始時点から第2のタイムアウト時間が経過したことを検出する第2の計時手段を前記第2の記憶手段上のキュー毎に具備し、該第2のタイムアウト時間の経過が検出された時点で該キューに滞留しているスーパーパケットを廃棄することを特徴としている。

【0035】

また、請求項8記載の発明は、請求項7記載の発明において、前記第2のタイムアウト時間は、予め決められた前記ネットワーク上のトラヒックの最小帯域または網内遅延の最大許容値、前記ネットワーク内の遅延分散時間、及び、予め定

められている保護時間から決定されることを特徴としている。

また、請求項 9 記載の発明は、請求項 1 ～ 8 の何れかの項記載の発明において、前記構築手段は、前記パケットが複数のスーパーパケットにまたがることを検出してこれら複数のスーパーパケットに分割して格納し、前記復元手段は、前記スーパーパケット上のパケットが複数のスーパーパケットにまたがっていることを検出し、これら複数のスーパーパケットに分割して格納されているパケットのデータを繋げて送信元のアクセス網が生成した元のパケットを復元することを特徴としている。

また、請求項 1 0 記載の発明は、請求項 1 ～ 9 の何れかの項記載の発明において、前記スーパーパケットには、そのペイロードが単一のパケットのデータで占有されているか否かを示す単一パケット占有情報が格納され、前記構築手段は、前記パケットの各々の長さを前記スーパーパケットのペイロードの長さと比較し、該比較結果に応じて前記単一パケット占有情報を占有有又は占有無に設定し、前記復元手段は、前記スーパーパケット上の前記単一パケット占有情報が占有有であることを検出し、前記単一パケット占有情報が占有無のスーパーパケットが到来するまでのこれら一連のスーパーパケット上に載せられているパケットのデータを繋げて送信元のアクセス網が生成した元のパケットを復元することを特徴とする。

【 0 0 3 6 】

また、請求項 1 1 記載の発明は、請求項 1 ～ 1 0 の何れかの項記載の発明において、前記スーパーパケットのヘッダは前記パケットのヘッダと同一のフォーマットを有していることを特徴としている。

また、請求項 1 2 記載の発明は、パケットを送受信するアクセス網と、請求項 1 ～ 1 1 の何れかの項記載のパケット転送装置と、パケットを中継する中継手段を備え、該中継手段のスイッチ単位である固定長セルの n 倍 (n は 2 以上の整数) の長さを持つスーパーパケットを転送するバックボーンのネットワークとを具備し、前記パケット転送装置は、前記アクセス網によって送受信される前記パケットと前記ネットワーク上で転送される前記スーパーパケットとを相互に変換して、送信元のアクセス網から送信される前記パケットを前記スーパーパケットの

形で前記ネットワーク内の前記中継手段を経由させて宛先のアクセス網まで転送することを特徴としている。

また、請求項 13 記載の発明は、請求項 12 記載の発明において、前記パケット転送装置及び前記中継手段には、前記ネットワーク内でのみ定義される固有の網内アドレスがそれぞれ付与されており、前記スーパーパケットのヘッダには、送信元アドレス及び宛先アドレスとして、送信元及び宛先のアクセス網にそれぞれ接続されたパケット転送装置の網内アドレスが格納されることを特徴としている。

また、請求項 14 記載の発明は、請求項 13 記載の発明において、宛先となるパケット転送装置が同一であるスーパーパケットのうち、該宛先のパケット転送装置に接続された宛先のアクセス網の種別に応じて、該スーパーパケットのヘッダ中の宛先アドレスに異なる網内アドレスが割り当てられていることを特徴としている。

【0037】

また、請求項 15 記載の発明は、請求項 13 又は 14 記載の発明において、前記中継手段は、前記スーパーパケットのヘッダ中の宛先アドレスと該スーパーパケットの中継先を対応づけたエントリが、前記パケット転送装置および前記中継手段の数に相当するだけ少なくとも格納された経路検索テーブルと、前記スーパーパケットのヘッダ中の宛先アドレスをもとに前記経路検索テーブルを検索して該スーパーパケットの中継先を特定する経路検索手段と具備することを特徴としている。

また、請求項 16 記載の発明は、請求項 12～15 の何れかの項記載の発明において、送信元のアクセス網に接続されたパケット転送装置から前記ネットワーク内の各中継手段を経て宛先のアクセス網に接続されたパケット転送装置に至るまでのトラヒックである網内フローが同一であれば、該同一の網内フローに対応するスーパーパケットのヘッダを互いに同じにしていることを特徴としている。

また、請求項 17 記載の発明は、送信元のアクセス網が、宛先のアクセス網に転送すべきパケットを生成して自身に接続された送信元パケット転送装置へ送信し、該送信元パケット転送装置は、前記送信元のアクセス網から送信されたパケ

ットをバックボーンとなるネットワーク上に設けられた中継手段のスイッチ単位である固定長セルの n 倍（ n は2以上の整数）の長さを持つスーパーパケットに変換して前記ネットワークに送出し、前記ネットワークは、前記中継手段を用いて前記スーパーパケットを中継して前記宛先のアクセス網に接続された宛先パケット転送装置に至るまで転送し、該宛先パケット転送装置は、前記ネットワークから転送される前記スーパーパケットをもとに前記送信元のアクセス網により生成された前記パケットを復元して前記宛先のアクセス網に送出することを特徴としている。

【0038】

また、請求項18記載の発明は、請求項17記載の発明において、前記送信元パケット転送装置は、前記送信されたパケットを前記宛先パケット転送装置を単位として個別に蓄積し、該宛先パケット転送装置毎に前記スーパーパケットを構築して前記ネットワークに送出し、前記宛先パケット転送装置は、前記ネットワークから転送されてくる前記スーパーパケットを前記送信元パケット転送装置を単位として個別に蓄積し、該送信元パケット転送装置毎に前記パケットの復元を行うことを特徴としている。

また、請求項19記載の発明は、請求項18記載の発明において、前記送信元パケット転送装置は、前記スーパーパケットが第1のタイムアウト時間にわたって構築されていないことを前記宛先パケット転送装置毎に検出し、該構築が行われていない宛先パケット転送装置に関して蓄積されているパケットが存在していれば、該パケットから前記スーパーパケットを構築して前記ネットワークに送出することを特徴としている。

また、請求項20記載の発明は、請求項18又は19記載の発明において、前記宛先パケット転送装置は、前記パケットが第2のタイムアウト時間にわたって復元されていないことを前記送信元パケット転送装置毎に検出し、該復元が行われていない送信元パケット転送装置に関して蓄積されているスーパーパケットが存在していれば、該スーパーパケットを廃棄することを特徴としている。

【0039】

また、請求項21記載の発明は、請求項17～20の何れかの項記載の発明に

において、前記送信元パケット転送装置は、前記送信元のアクセス網から送信された前記パケットが複数のスーパーパケットにまたがる場合に該パケットをこれら複数のスーパーパケットに分割して前記ネットワークに送出し、前記宛先パケット転送装置は、前記ネットワークから転送されてくる前記スーパーパケット中のパケットが複数のスーパーパケットにまたがっている場合に、これら複数のスーパーパケットに分割されているパケットのデータを繋げて前記送信元のアクセス網によって生成された元のパケットを復元して前記宛先のアクセス網に送出することを特徴としている。

また、請求項 2 2 記載の発明は、請求項 1 7 ~ 2 1 の何れかの項記載の発明において、前記送信元パケット転送装置は、前記スーパーパケットのヘッダ中の送信元アドレス及び宛先アドレスとして、該送信元パケット転送装置及び前記宛先パケット転送装置にそれぞれ付与されている前記ネットワーク内でのみ定義される固有の網内アドレスを格納して前記ネットワークに送出し、前記ネットワーク内の各中継手段は、転送されてくる前記スーパーパケットのヘッダを参照して前記宛先パケット転送装置に付与された網内アドレスに従って該スーパーパケットの中継先を特定して該スーパーパケットを前記宛先パケット転送装置まで転送することを特徴としている。

【 0 0 4 0 】

【発明の実施の形態】

以下、図面を参照して本発明の一実施形態について説明する。本発明はネットワーク上で転送されるパケットやフレーム一般に適用することが可能であるが、本実施形態では、こうしたパケットやフレームのうち、IP パケットを対象としたルータ装置を例に挙げて説明するものとする。

【 0 0 4 1 】

(1) ネットワークの構成

図 1 は本実施形態が適用されるパケット通信システムの構成を示したブロック図である。同図に示したように、本実施形態を実現するためのネットワークはコアルータ装置、エッジルータ装置という 2 種類のルータ装置を具備していることを特徴の一つとしている。本実施形態のネットワークではこれらルータ装置以外

にも、図15で説明した従来のネットワークと同様のユーザLANが配置されている。そして、コアルータ装置同士、エッジルータ装置同士、コアルータ装置とエッジルータ装置の間がそれぞれ伝送路100で接続されており、エッジルータ装置とユーザLANとの間がアクセス回線で接続されている。なお、繁雑になることから同図では敢えて伝送路の全てについて符合を付けることはしていない。

【0042】

さらに詳述すると、同図では、伝送路100を介してループ状に接続された8台のエッジルータ装置101a~101hと、伝送路100を介して相互に接続された4台のコアルータ装置102a~102dと、アクセス回線103b, 103c, 103d, 103f, 103gを介してエッジルータ装置101b, 101c, 101d, 101f, 101gにそれぞれ接続されたユーザLAN104b, 104c, 104d, 104f, 104gを例示してある。また、各コアルータ装置はそれぞれ伝送路100を介して隣接する2つのエッジルータ装置に接続されており、例えばコアルータ装置102aはエッジルータ装置101a, 101bに接続されている。

【0043】

このほか、各ユーザLANには図15と同様にホスト端末が接続されており、同図ではユーザLAN104cに接続された送信元ホスト105cとユーザLAN104gに接続された宛先ホスト105gだけを示してある。なお、エッジルータ装置、コアルータ装置、ユーザLAN、ホスト端末の台数は任意数であって良く、エッジルータ装置に接続される他のエッジルータ装置やコアルータ装置やユーザLANの台数、ユーザLANに接続されるホスト端末の台数、コアルータ装置に接続される他のコアルータ装置やエッジルータ装置の台数なども任意であって良いのは勿論である。

【0044】

例えば、ユーザLANに接続されているアクセス回線の回線速度は遅く、例えばイーサネットであれば10M~100Mbps程度である。一方、エッジルータ装置やコアルータ装置はT（テラ）bps級の能力を備えており、これらを接続する個々の伝送路100の速度も例えば数G~10Gbps程度に達する。し

たがって、例えば1台のエッジルータ装置に10Gbpsの伝送路が10本接続されていることを想定すると、100Mbpsのアクセス回線であれば千本単位の回線が1台のエッジルータ装置に接続されることになる。

【0045】

また、図1ではエッジルータ装置にユーザLANが接続されている形態を例示してあるが、こうしたユーザLAN以外にも既存の一般的なアクセス網、既存の交換機、既存のルータ装置、SDH（同期デジタルハイアラキ）などで使用される既存の通信網装置などであっても良い。図2はこうしたアクセス網を例示したものであって、エッジルータ装置101とユーザLAN120との間がイーサネットやフレームリレー網を介して接続されている形態、エッジルータ装置101とユーザLAN121との間が既存のアクセスルータ装置122を介して接続されている形態、エッジルータ装置101とATM（非同期転送モード）回線の間が既存のATMアクセス装置123で接続されている形態、エッジルータ装置101とSTM（同期転送モード）回線との間が既存のSTM装置124で接続されている形態が混在している場合を示してある。

【0046】

図1を再び参照すると、本実施形態では、エッジルータ装置101a～101hおよびこれらを接続している伝送路100から形成される枠の内側がパケット通信システムのバックボーンとなるネットワークを構成しており、以下ではこれをネットワーク106と定義する。つまり、本実施形態のパケット通信システムではエッジルータ装置101a～101hを境にしてネットワーク106と各ユーザLANとの間が区切られていると考えることができる。このネットワーク106内では当該ネットワーク内でのみ使用される独自のアドレス体系たる「網内アドレス」を用いて運用されている。

【0047】

図中、括弧でくくられた数字“1”～“12”がエッジルータ装置又はコアルータ装置にそれぞれ付与された固有の網内アドレスを示しており、例えばエッジルータ装置101aには網内アドレスとして“1”が付与されている。網内アドレスの値はネットワーク106を構成するエッジルータ装置およびコアルータ装

置に対してそれぞれ一意に設定されており、ネットワーク 106 内では網内アドレスのみを手がかりにパケット（後述するスーパーパケット 110）を転送することが可能である。換言すれば、個々のユーザ LAN から送出される IP パケットのヘッダに含まれている宛先 IP アドレスがネットワーク 106 内のコアルータ装置で参照されることはない。

【0048】

また、本実施形態ではネットワーク 106 内のトラヒックの流れを網内フローと呼んでいる。図 1 に例示されている網内フロー 107 はネットワーク 106 の入口（「Ingress」）に存在するエッジルータ装置 101c からネットワーク 106 の出口（「Egress」）に存在するエッジルータ装置 101g に転送されるパケットのトラヒックとして規定される。なお以下では、網内フロー 107 上でエッジルータ装置 101c に相当するエッジルータ装置を「送信元エッジルータ」または「Ingress ルータ」と呼ぶことがあり、また、網内フロー 107 上でエッジルータ装置 101g に相当するエッジルータ装置を「宛先エッジルータ」または「Egress ルータ」と呼ぶ場合がある。

【0049】

次に、各コアルータ装置はネットワーク 106 内に設けられた所定の中継地点に配備されている。これらコアルータ装置は、隣接するエッジルータ装置ないしは他のコアルータ装置から転送されてくる固定長のスーパーパケットを単位にルーティングを行って転送するルータ装置である。このため、本実施形態のコアルータ装置では、従来技術で説明したルータ装置のように、パケットないしフレームを装置内セルに分割して転送する必要がなく、スーパーパケットを単位としてネットワーク 106 内をそのまま転送させることが可能である。

【0050】

ここで、後述するようにスーパーパケットはネットワーク 106 内でのみ定義される網内フレームであって、通常の IP パケットと同形式のヘッダを有していることから、IP レイヤ上ではスーパーパケットが通常の IP パケットと同じように扱われる。こうしたことから、コアルータ装置は通常用いられている IP ルータ装置で構成することが可能である。しかも、コアルータ装置の転送対象であ

るスーパーパケットのヘッダには網内アドレスの形式で表された宛先アドレスが書き込まれていることから、次のような利点もある。

【0051】

すなわち、一般的なIPルータ装置はあらゆる宛先IPアドレス（例えば32ビットで表現される数のアドレス）に対応したアドレス解決テーブル（図16の経路検索テーブル17に相当）を備えている。これに対して、コアルータ装置はネットワーク106内で定義される比較的少数の網内アドレスに対応した数のエントリを持つアドレス解決テーブルを構成すれば済む。このため、例えば図1ではコアルータ装置内のアドレス解決テーブル（経路検索テーブル）が少なくとも12個のエントリを有していれば良くなり、アドレス解決テーブルを従来に比べて小規模なものにすることができる。このほか、スーパーパケットの長さを大きくすればするほどコアルータ装置におけるスイッチ回数が減少するため、それだけコアルータ装置をより大規模化することも可能である。

【0052】

一方、各エッジルータ装置はネットワーク106の境界に位置するルータ装置であって、ネットワーク106と各ユーザLANとの間でデータ転送を行う際の変換処理を行う。こうしたことから、本実施形態ではユーザLANなどの既存のアクセス網とネットワーク106との間には最低限エッジルータ装置を配する必要があるものの、例えば図1においてエッジルータ装置101aとエッジルータ装置101bを接続する伝送路上にコアルータ装置を配置するような構成としても良い。

【0053】

次に、本実施形態では、ユーザLAN側において図17で説明したIPパケット37（正確にはIPパケット37が載せられたレイヤ1フレーム）を単位とした転送が行われる。もっとも、ネットワーク106内では次に述べるように共通フレームであるスーパーパケットに従って転送が行われるため、アクセス網側における転送がIPパケットを単位とした転送に限られるものではない。例えば、図2について説明したようにATM、音声、STMなどのトラヒックをスーパーパケットに寄せ替えて収容することも可能であるが、以下では説明の都合からI

P パケットを例に挙げて説明することにする。

【0054】

その一方で、ネットワーク 106 内では IP パケット単位の転送ではなく、図 17 に示した装置内セル 40 の長さを L としたとき、この固定長 L の整数倍の長さを持ったスーパーパケット 110 を単位として転送が行われる。すなわち、 n を 2 以上の整数とすればスーパーパケット 110 は固定長 “ $n \times L$ ” の大きさを有していることになる。なお、 n の値が “1” の場合は装置内セルそのものであるため n を “2” 以上としてある。また、この固定長 “ $n \times L$ ” をこれ以後の説明中で「スーパーパケット長」と呼ぶことがある。このスーパーパケット 110 はネットワーク 106 の内部でのみ定義されるパケットであって、換言すればネットワーク 106 内ではスーパーパケット 110 のみが転送されることになる。

【0055】

スーパーパケット長は通常 1500 バイト～2000 バイト程度にしてあり、個々の IP パケットの平均的な長さに対して $1/10$ 程度となっている。もっとも、本実施形態を実現するにあたり、スーパーパケット長には上限は全くなくネットワークを構築する際にその大きさを適宜決定すれば良い。

なお、従来技術のところで述べたように、IP パケット 37 やスーパーパケット 110 は何れもレイヤ 1 フレームの形でアクセス回線上または伝送路 100 上を伝送することになるが、煩雑さを避けるために、「IP パケット 37 やスーパーパケット 110 が伝送路やアクセス回線上を転送される」といった表現を用いることがある。

【0056】

以下に述べるようにスーパーパケット 110 自体は IP パケット 37 と同じフレームフォーマット（パケットフォーマット）を有しているため、上述した通り、各コアルータ装置は通常の IP ルータ装置で実現できる。一方のエッジルータ装置であるが、例えば送信元エッジルータ 101c は、宛先エッジルータ 101g を含めた各宛先エッジルータを単位として、ユーザ LAN 104c から到来する様々な長さの IP パケットをスーパーパケット 110 のフレームフォーマットに変換している。これとは逆に、宛先エッジルータ 101g はネットワーク 10

6上の網内フロー107を通じて転送されてくるスーパーパケット110から元のIPパケットを取り出す処理を行っている。こうしたスーパーパケットの組立機能・解体機能は通常のルータ装置には見られない機能であって、本実施形態によるエッジルータ装置に独特の機能である。

【0057】

(2) スーパーパケットのフレーム構成

以下にスーパーパケットの詳細なフォーマットについて説明する。図3に示したように、スーパーパケット110のフレームフォーマットは前述した通常のIPパケットと同様のフレームフォーマットとなっている。すなわち、スーパーパケット110は最初にパケットヘッダ111が定義され、これに続いてパケットペイロード112が定義されている点においてIPパケットと共通している。また、図4に示したようにスーパーパケット110のパケットヘッダ111のフォーマットもIPパケットのヘッダフォーマットと同じになっている。

【0058】

その一方で、IPパケットはその全体の長さが可変長であるのに対し、スーパーパケット長は、図15ないし図16に示した従来のルータ装置におけるスイッチ単位である装置内セルの長さの整数倍の長さを持った固定長のパケットになっている点でIPパケットと相違している。また、スーパーパケット110は、ネットワーク106内の送信元エッジルータ（Ingressルータ）101cから同ネットワーク内の宛先エッジルータ（Egressルータ）101gに至るトラヒック転送である網内フロー107に対応するものであって、パケットヘッダ111で区別されるスーパーパケット110の種別と網内フロー107とが1:1に対応している。このため、網内フロー107の経路上におけるパケットヘッダ111は網内フロー107上を通じて常に同一のフォーマットになる。

【0059】

ここで、図4を参照してスーパーパケット110のヘッダフォーマットについて説明する。なお、同図ではIPv4パケットと同様のヘッダフォーマットを採用した場合について示してあるが、例えばIPv6パケットを使用した場合もスーパーパケットのヘッダフォーマットが通常のIPパケットと同じであることに

変わりはない。もっとも、ヘッダの各フィールドに格納される値には両者の間で若干の差異がある。すなわち、IP パケットではヘッダ部分の送信元アドレス及び宛先アドレスのフィールドには送信元 IP アドレスおよび宛先 IP アドレスが格納される。

【0060】

これに対して、スーパーパケット 110 のパケットヘッダ 111 では、送信元アドレス及び宛先アドレスの各フィールドに対し、それぞれネットワーク 106 内の送信元および宛先に相当するエッジルータ装置に付与された網内アドレス（即ち、送信元エッジルータアドレスおよび宛先エッジルータアドレス）が格納される。例えば図 1 に示した網内フロー 107 の場合、パケットヘッダ 111 内の送信元アドレスとしてエッジルータ装置 101 c のアドレス値“3”が記述され、パケットヘッダ 111 内の宛先アドレスとしてはエッジルータ装置 101 g のアドレス値“7”が記述されることになる。

【0061】

なお、送信元アドレス及び宛先アドレス以外のフィールドについても以下に説明しておく。まず、「バージョン」はどのバージョンの IP ソフトウェアによって送信されたパケットかを表しており、この場合はバージョン 4 を示すデータが格納される。「IHL」は Internet Header Length の略称であってパケットヘッダ 111 の大きさを 4 オクテット単位で表したものである。「サービスタイプ」は送信している IP サービスの種類を示すものである。「トータル長」はパケットヘッダ 111 およびパケットペイロード 112 からなるスーパーパケット 110 全体の長さをオクテット単位で表したものである。「識別子 (ID)」は上位層（後述）に対してデータを渡すときの参考情報としてパケット毎に固有の番号が設定されるフィールドである。「フラグ」はパケットの分割に関する情報を示しており、分割の可否や分割されたパケットであればそれが最後のパケットか否かなどを表している。

【0062】

「フラグメント・オフセット」はパケットが分割されている場合に、元のデータにおける先頭からのオフセットをオクテット単位で表したものである。「生存

期間 (T T L)」はTime To Liveの略称であって、当該パケットがネットワークに存在を許されている生存期間を秒単位で表したものである。「プロトコル」は上位層として用いられるプロトコルを表しており、ICMP (Internet Control Message Protocol) , TCP (Transmission Control Protocol) , UDP (User Datagram Protocol) 等である。「ヘッダチェックサム」はパケットヘッダの内容の正当性を確認するためのチェックサムである。「オプション」は標準のパケットヘッダに存在しない情報を送信するためのフィールドである。「パディング」はオプションフィールド全体を4オクテット単位にするための空き領域である。

【0063】

一方、再び図3を参照すると、スーパーパケット110のパケットペイロード112には図1に示した各ユーザLANから送出される1ないし複数のIPパケットが詰め込まれている。このほか、1つのIPパケットの一部でパケットペイロード112の全体が埋められる場合もあるがこれについては後述する。そして、例えば図1に示した送信元エッジルータ101cは、宛先エッジルータ101gに向かう網内フロー107等の網内フローを単位として、ユーザLAN104cから到着する可変長のIPパケットを集約し、これらIPパケットを網内フロー107単位に分類されたスーパーパケット110のパケットペイロード112部分に詰め込んでネットワーク106に転送する処理を行う。

【0064】

こうしたことから、最後に詰め込まれるIPパケットがパケットペイロード112に入り切らなくなると、1個のIPパケットが複数のスーパーパケットに分割される状況も発生しうる。こうした状況を図5(a)に例示してあり、同図では順次入力される一連のパケットPA~PHをスーパーパケットに詰め込む場合についてのものである。ここで、スーパーパケットA1/A2は同一の網内フロー107に対応して順次転送されるパケットであることから、これらスーパーパケットは互いに同じパケットヘッダ111を有している。このうち、スーパーパケットA1は図3に示したスーパーパケットと全く同じものである。

【0065】

いま、可変長のIPパケットをスーパーパケット110に詰めていくとそれら

IP パケットを集めたものも当然ながら可変長になる。そしていま“ $L1 + L2$ ”の大きさを持つパケットPFをスーパーパケットA1に挿入しようとする、パケットペイロード112に残されている領域（長さ $L1$ ）よりもパケットPFが大きいため、パケットPFの先頭から長さ $L1$ だけのデータしか挿入することができない。このため、残っている長さ $L2$ のデータは、次のスーパーパケットA2を送出する時にこのスーパーパケットのパケットペイロード112の先頭に挿入され、これに引き続いて新たなパケットPG, PHが順次スーパーパケットA2に挿入されてゆく。なお、パケットPAもパケットPFと同様に、先行するスーパーパケット（図示省略）とスーパーパケットA1とにまたがって転送されることを想定している。

【0066】

このようにして、パケットPA～PE及びパケットPFの一部がスーパーパケットA1に詰め込まれ、パケットPFの残部及びパケットPG, PHがスーパーパケットA2に詰め込まれる。かかる場合に対応するために、スーパーパケット110のパケットペイロード112の先頭には、図3に示すように、ポインタフィールド（図中の①）、パケット数フィールド（図中の②）、Padding有無フィールド（図中の③）、単一パケット占有フィールド（図中の④）がそれぞれ定義されている。これらフィールドを設けることによって、可変長のIPパケットを固定長のスーパーパケットのパケットペイロード部分に順番に詰めてゆくことが可能になる。

【0067】

ここで、ポインタフィールド①はパケットペイロード112に現れる最初のパケットの先頭位置を示したポインタである。上述したように、図3に示したパケットPAは先行するスーパーパケットとの間にまたがっているため、同図に示したパケットPAについてはパケットの先頭位置を観念することができない。このため、パケットPBが図3のスーパーパケット110上では先頭のパケットとなって、ポインタフィールド①がパケットPBの先頭を指している。

【0068】

なお、パケットPAが複数のスーパーパケットにまたがらない独立したパケッ

トであれば、ポインタフィールド①はこのパケット P A の先頭位置を指し示すことになり、ポインタフィールド①に格納される値は“0”となる。つまり、ポインタフィールド①の値はパケットヘッダ 1 1 1 の直後を基準位置とした値になっている。したがって、ポインタフィールド①が“0”であれば図 3 のパケット P A に相当するパケットは複数のスーパーパケットにまたがっておらず、これとは反対に当該フィールドの値が“0”以外であれば複数のスーパーパケットにまたがっていることが判別できる。なお、基準位置を示すポインタ値は“0”でなくとも良く例えば“1”などであっても良い。

【0069】

一方、パケット P B の次に詰め込まれるパケット P C の先頭位置は、パケット P B のヘッダに記述されたトータル長を新たなポインタの代わりに使用することで算出可能である。すなわち、I P パケットのヘッダは図 4 に示したスーパーパケット 1 1 0 のパケットヘッダ 1 1 1 と同一フォーマットであるため、パケット P B の先頭位置を示すポインタフィールド①に対して当該パケット P B のトータル長を加算することでパケット P C の先頭位置が得られる。これは、パケット P C に続くパケット P D, P E も同じであって、パケット P C と同様の手順に従ってパケット P C, P D のトータル長をさらに加算してゆくことでそれぞれのパケットの先頭位置を特定することが可能である。

【0070】

次に、パケット数フィールド②は、パケットペイロード 1 1 2 の最後に収容されているパケットの先頭位置を特定するためなどの目的で設けられたものであって、パケットペイロード 1 1 2 内に現れる I P パケットの先頭位置の個数を示している。そして、最後に収容されているパケットの先頭位置を特定するには、上述した過程、すなわち、各 I P パケットのヘッダに含まれるトータル長を擬似ポインタと見なしてポインタフィールド①で指定される位置から I P パケットを順に辿ってそれらの先頭位置を特定してゆく過程において、これと並行して、パケットペイロード 1 1 2 に詰め込まれている I P パケットの発生回数をカウントしてゆけば良い。

【0071】

そしてもし「発生回数<パケット数フィールド②の値」であればその時点で着目しているIPパケットに後続するIPパケットがさらに存在しているので、現時点におけるIPパケットのヘッダに含まれるトータル長を利用して次のIPパケットの先頭位置を算出する。これに対して、「発生回数=パケット数フィールド②の値」となっていれば、現時点で着目しているIPパケットがパケットペイロード112に詰め込まれている最後のIPパケットの領域であるものと認識する。例えば図5(a)の場合、スーパーパケットA1上にはパケットPAの先頭位置が存在せず、パケットPA/PB間、PB/PC間、PC/PD間、PD/PE間、PE/PF間の5箇所にIPパケットの先頭位置があるので、パケット数フィールド②の値としては“5”が設定される。したがって、上記過程の中で5番目に認識されるパケットPE/PF間の境界が最後に収容されているパケットの先頭位置となる。

【0072】

次に、Padding有無フィールド③は以下のような理由で設けられている。すなわち、上述したポインタフィールド①だけであると、図5(a)に示されるパケットPFの領域がスーパーパケットA1/A2に分割されて実際にIPパケットデータが格納された領域であるのか、あるいは、図5(b)で示されているように単なる空き領域としてPaddingされているのか判別することができない。そこで、パケットPEの後ろにあるパケットペイロード112上の領域が新たなパケット(パケットPF)の一部であるかそれとも空き領域としてPaddingされているかのPadding有無フィールド③の内容に従って判定するようにしている。

【0073】

例えば図5(b)の場合には当該領域が空き領域となっているため、スーパーパケットA1のPadding有無フィールド③は「Padding有」を示すデータが格納される。一方、図5(a)の場合、当該領域には新しいパケットであるパケットPFの先頭から長さL1に相当するだけのデータが格納されており、長さL2に相当するパケットPFの残りのデータは同じ網内フロー107(図1参照)で転送されるスーパーパケットA2の先頭に格納される。したがってこの場合、スーパーパケットA1のPadding有無フィールド③に「Padding無」を示すデータが格納

される。

【0074】

またこのとき、次のスーパーパケットA2については、残りのデータ長L2をポインタフィールド①に書き込むようにしている。これによって、パケットPFの次のIPパケットであるパケットPGの先頭位置を指し示すことができる。このようにすることで、スーパーパケットA1/A2を順次受け取った宛先エッジルータ101gでは、スーパーパケットA2の先頭に詰め込まれている長さL2の領域と、スーパーパケットA1の最後に収容されているパケットPFの先頭位置から長さL1分の領域とを繋げることによって、元のIPパケットであるパケットPFを復元することが可能となる。

【0075】

なお、IPパケットをスーパーパケットに詰め込んでゆく過程において、例えば図5(a)に示したパケットPFがスーパーパケットA1、A2にまたがるのか、或いは、パケットPFがスーパーパケットA1にちょうど収まるのかも容易に判別できる。すなわち、スーパーパケット長は予め決められた固定値であるため、スーパーパケットのパケットペイロードからIPパケットを取り出す度にパケットペイロードの残りの長さを求めてゆき、最後に格納されているパケットPFのヘッダ中のトータル長を参照してこれがパケットペイロードの残りの長さ(図5(b)の空き領域に相当する長さ)よりも大きければ、当該IPパケットはスーパーパケットA1、A2にまたがることが分かり、一方、両者が一致していればパケットPFはスーパーパケットA1にちょうど収まることが分かる。

【0076】

また、Egressルータ側でスーパーパケットA1からパケットPFを取り出す場合にも、当該パケットがスーパーパケットA1、A2にまたっているのかどうかは容易に判断できる。すなわち、スーパーパケットA1から取り出したパケットPFの長さと当該パケットのヘッダ中に記載されているトータル長を比較し、両者が一致していればパケットPFは複数のスーパーパケットにまたがっておらず、一方、両者が不一致していればパケットPFは複数のスーパーパケットにまたがっていることが分かる。

【0077】

また、異なる網内フローに対応するスーパーパケット間では、上述したようなスーパーパケットにまたがるIPパケットのデータ割当を実施することはできず、同一の網内フローにおいてのみ可能である。そして、同一の網内フロー上を転送されるスーパーパケットのパケットヘッダは互いに同じであるから、スーパーパケット間でIPパケットを繋げる処理は同一のパケットヘッダを有するものを対象として行うことになる。

【0078】

なお、Padding有無フィールド③を設けることなく、IPパケットが複数のスーパーパケットにまたがっているか或いはPaddingであるのかを判別することも可能である。例えば、図5(a)は複数のスーパーパケットにまたがる場合であって、スーパーパケットA1のパケット数フィールド②の値は“5”であり、当該フィールドで指定される最後のパケットはパケットPFである。一方、図5(b)はPaddingが為される場合であり、スーパーパケットA1のパケット数フィールド②の値は“4”であって、当該フィールドで指定される最後のパケットはパケットPEである。そこで、最後のパケットのヘッダに記載されているトータル長に基づいて当該最後パケットの最終位置の直後の位置を算出すると、図5(a)の場合にはスーパーパケットA1の範囲外となり、一方で、図5(b)の場合には空き領域の先頭位置となる。こうした違いから、前者のようになれば複数のスーパーパケットにまたがっているものと判定することができ、後者のようになればPaddingされていると判定できる。

【0079】

次に、単一パケット占有フィールド④は以下のような理由で設けられている。図5に例示した場合のように、各IPパケットがスーパーパケット長に比べて短く、複数のIPパケットが単一のスーパーパケット110のパケットペイロード112に詰め込まれる場合には上述した①～③のフィールドを設けることで足りる。これに対して、図6で示すように単一のIPパケット37の大きさが数10キロバイト程度と長く、IPパケット37が複数のスーパーパケットA1～A7にまたがって転送される場合も想定しうる。こうした場合、あるスーパーパケッ

トの packets ペイロード 112 の全体が同一の IP パケットのデータだけで占有されることになって、当該スーパーパケットには IP パケットの先頭も空き領域 (Padding) もなくなる場合がある。

【0080】

図 6 の場合で言えば、スーパーパケット A1～A6 はいずれも同一の IP パケット 37 のデータによって占有されており、これらスーパーパケットに関してポインタフィールド①、パケット数フィールド②、Padding 有無フィールド③はいずれも意味を為さなくなる。そこでこうした場合のために、単一パケット占有フィールド④が“ON”のときにはパケットペイロード 112 が一つの IP パケットのデータで占有されているものと判断し、その一方で、当該フィールドが“OFF”のときには複数の IP パケットがパケットペイロード 112 内に詰め込まれているものと判断する。

【0081】

そのために、未だスーパーパケットに詰め込まれていない IP パケットの残りの長さ、スーパーパケット 110 のパケットペイロード 112 の長さを比較して、前者の長さが後者の長さ以上であれば単一パケット占有フィールド④を“ON”に設定し、さもなければ“OFF”に設定する。したがって、図 6 に示したスーパーパケット A1～A6 を転送する場合には単一パケット占有フィールド④の値を“ON”に設定する一方で、図 6 に示したスーパーパケット A7 を転送する場合や図 5 に示した一連の IP パケットを転送する場合には単一パケット占有フィールド④の値を“OFF”に設定することになる。

【0082】

ここで、同一の IP パケットは宛先エッジルータアドレスが同じ 1 台のエッジルータ装置に転送される。このため、或る IP パケットが複数のスーパーパケットに分割されて転送される場合、これら複数のスーパーパケットでは同じ網内フロー 107 に対応した同一のパケットヘッダ 111 が付与されて、ネットワーク 106 内の同じ経路（例えば、送信元エッジルータ 101c→コアルータ装置 102b→コアルータ装置 102d→宛先エッジルータ 101g）上で転送される。こうしたことから、宛先エッジルータ 101g が、同じ送信元エッジルータア

ドレスを有するスーパーパケットを同一のキューで管理することによって、複数のスーパーパケットに分割された1個のIPパケットを元のIPパケットに復元することが可能になる。

【0083】

最後に、本実施形態の特徴であるスーパーパケットが通常のIPパケットと異なる点を列挙しておく以下のようなになる。

- a) パケット長が固定長である。
- b) パケットペイロード112に複数のIPパケットを挿入可能である。
- c) 一つのIPパケットが複数のスーパーパケットにまたがって挿入されることを考慮して、パケットペイロード112上にポインタフィールド①が定義されている。
- d) パケットペイロード112上に現れるIPパケットの先頭位置の数がパケットペイロード112上のパケット数フィールド②に定義されている。
- e) パケットペイロード112に含まれている最後のパケットの領域が空き領域としてPaddingされているか否かを示すPadding有無フィールド③がパケットペイロード112上に定義されている。
- f) 一つのIPパケットがスーパーパケット長よりも大きい場合を考慮に入れて、パケットペイロード112が単一のIPパケットの部分データで占有されているか否かを示す単一パケット占有フィールド④がパケットペイロード112上に定義されている。

【0084】

(3) エッジルータ装置の構成

次に、図1に示した各エッジルータ装置の構成について図7に示したブロック図を参照して説明する。図7に示されるエッジルータ装置201は図1の各エッジルータ装置に対応するものであって、アクセスインタフェース（図中の「INF」）部202-1～202-M（M：自然数），スイッチ部203，スーパーパケット処理部204-1～204-P（P：自然数）に大別される。なお、“M”の値および“P”の値はいずれも任意数であって、ネットワーク構成やルータ装置の規模によって適宜決定すれば良い。そして本実施形態のエッジルータ装置201

は、ネットワーク 106 側のインタフェースとしてスーパーパケット処理部 204-1~204-P を有する点を除けば、機能的には従来技術で説明したルータ装置と同じである。

【0085】

すなわち、スイッチ部 203 は図 16 に示したスイッチスケジューラ部 12 およびクロスバースイッチ部 13 が具備する機能をそれぞれ合わせたものと等価であって、従来技術で説明したルータ装置と同様に、固定長 L を持った装置内セルを単位としてスイッチング動作を行うものである。一方、アクセスインタフェース部 202-1~202-M は、いずれも、図 16 に示した各入力インタフェース部および各出力インタフェース部がそれぞれ具備する機能を併せ持っている。これらアクセスインタフェース部の各々は図 1 に示したアクセス回線を介してユーザ LAN に接続している。

【0086】

各アクセスインタフェース部は、図 16 に示した入力インターフェイス部と同様に、当該ユーザ LAN からアクセス回線に送出されるレイヤ 1 フレームから IP パケット 37 を取り出し、当該 IP パケット 37 のヘッダ部分に記述されている宛先 IP アドレスに基づいて何れのネットワーク回線へ IP パケットを送出すべきかを決定する。そして、各アクセスインタフェース部は、取り出された IP パケット 37 を固定長の装置内セル 40 に分割してスイッチ部 203 へ送出するとともに、装置内セル 40 のセルヘッダ 41 に書き込んでおいた出力インタフェース番号に相当するスーパーパケット処理部 204-1~204-P の何れに転送すべきかをスイッチ部 203 に通知する。こうしてアクセスインタフェース部は、スイッチ部 203 を通じて、決定されたネットワーク回線を終端しているスーパーパケット処理部 204-1~204-P の何れかへ当該 IP パケットを転送する。

【0087】

また以上とは逆に、各アクセスインタフェース部は、スーパーパケット処理部 204-1~204-P の何れかのスーパーパケット処理部から送られてくる IP パケットを装置内セルに分割された形でスイッチ部 203 から受け取る。そして各アクセスインタフェース部は、図 16 に示した出力インタフェース部と同様に、

装置内セルからIPパケットを復元してこれにIPレイヤ処理を施したのち、レイヤ1/2のヘッダを付与することでレイヤ1フレームを構築し、これをアクセス回線経由でユーザLAN側に出す。

【0088】

(a) スーパーパケット処理部の構成

本実施形態の特徴であるスーパーパケット処理部204-1~204-Pは、図1に示したユーザLAN側からネットワーク106側への転送を行うスーパーパケット構築部205と、ネットワーク106側からユーザLAN側への転送を行うスーパーパケット復元部206とで構成されている。そのために、スーパーパケット構築部205は装置内セルより抽出したIPパケットからスーパーパケットへの変換を行ってスーパーパケットが含まれたレイヤ1フレームを構築する機能を有している。一方で、スーパーパケット復元部206はレイヤ1フレームより抽出されるスーパーパケットからIPパケットへの変換を行って元のIPパケットを復元してこれを装置内セルの形で送る機能を有している。

【0089】

(b) スーパーパケット構築部の構成

次に、図8を参照してスーパーパケット構築部205の詳細について説明する。図示したように、スーパーパケット構築部205は、従来技術で説明したものと同一パケット復元部18(図16)、宛先エッジルータ判定部301、書き込み制御部302、読み出し制御部303、バッファ部304、タイマー部305、ポインタ部306、レイヤ1/レイヤ2(図中では「L1/L2」)処理部307、アドレス解決テーブル308より構成されている。そして、パケット復元部18は図7に示したスイッチ部203に接続されており、レイヤ1/レイヤ2処理部307は図1に示した伝送路100に接続されている。

【0090】

まず、パケット復元部18は、図7に示したアクセスインタフェース部およびスイッチ部203を介して送られてくる固定長の装置内セルを積み上げることで、アクセス回線に送られた元々のIPパケットを復元し、これを宛先エッジルータ判定部301に転送する。次に、宛先エッジルータ判定部301は、パケッ

ト復元部 18 から個々の IP パケットが入力されると、各 IP パケットのヘッダに含まれている宛先 IP アドレスを元にアドレス解決テーブル 308 を検索して宛先エッジルータアドレスを解決し、この宛先エッジルータアドレスを当該 IP パケットと一緒にして書き込み制御部 302 に転送する。

【0091】

ここで、宛先エッジルータ判定部 301 が行うアドレス解決処理とアドレス解決テーブル 308 のテーブルフォーマットについて説明する。アドレス解決テーブル 308 のフォーマットは図 9 に示される通りであって、Patricia Trie と呼ばれる Binary Tree 状のデータベースによって構成されている。Binary Tree の構造は一般的な技術であって、通常の IP ルータ装置にも搭載されている機能であるためここではその詳細は割愛するが、例えば IP アドレスに関する Binary Tree の参考文献として、「Radix: Keith Sklower, "A Tree-Based Packet Routing Table for Berkeley Unix", USENIX Winter, pp.93-99, 1991」がある。

【0092】

アドレス解決テーブル 308 の各エントリは、宛先 IP アドレス（この場合は IP v4 であるため 32 ビット）、当該宛先 IP アドレスに対応する Egress エッジルータに付与されている網内アドレス、Binary Tree 上で隣接している節点ないし葉に対応したエントリとの間の接続を示すポインタ情報から構成されている。このアドレス解決テーブル 308 はあらゆる宛先 IP アドレスに対応してそのエントリが構築される。そして宛先エッジルータ判定部 301 は、ユーザ LAN から転送されてくる IP パケットのヘッダに書き込まれている宛先 IP アドレスをもとにアドレス解決テーブル 308 上で検索を行って、当該 IP パケットが最終的に転送されるところの Egress エッジルータに付与される網内アドレスを求めて、これを宛先エッジルータアドレスとして書き込み制御部 302 に送出する。

【0093】

次に、書き込み制御部 302 は、図 10 に示したように、バッファ部 304 を利用してネットワーク 106 内の宛先エッジルータに付与された網内アドレスを単位としたキューを管理しており、これらキューのうち、宛先エッジルータ判定部 301 から通知された宛先エッジルータアドレスに対応するキューに対して、

同じく宛先エッジルータ判定部 301 から転送される IP パケットを格納する。
また、書き込み制御部 302 は、バッファ部 304 に書き込んだ IP パケットに関する各種の情報（即ち、当該パケットのヘッダに含まれているトータル長など）を読み出し制御部 303 に通知する。

【0094】

次に、読み出し制御部 303 は、書き込み制御部 302 から通知された各種の情報に基づいてバッファ部 304 内に存在する各キューに蓄積されているデータ長（キュー長）を管理している。そして読み出し制御部 303 は、各キューに蓄積された IP パケットの総データ長がスーパーパケット長に達していることを検出し、当該条件に該当するキューから蓄積されている IP パケットを読み出す。もっとも、正確に言うと、図 3 に示したようにスーパーパケット 110 にはパケットヘッダ 111 が含まれるほか、パケットペイロード 112 には①～④のフィールドが設けられている。そこで実際には、これらに必要となる長さをスーパーパケット長“ $n \times L$ ”から減じた分のデータが蓄積された時点でキュー上から IP パケットを読み出すことになる。そして読み出し制御部 303 は、読み出された 1 ないし複数の IP パケットに対してパケットヘッダ 111 を付与するとともに、図 3 に示した①～④の各フィールドの領域に対して IP パケットの状況に応じたデータを書き込んでスーパーパケット 110 を構築してレイヤ 1 / レイヤ 2 処理部 307 に転送する。

【0095】

次に、バッファ部 304 は上述したように宛先エッジルータ判定部 301 から転送される IP パケットを蓄積するためのメモリであって、同じく宛先エッジルータ判定部 301 から通知される宛先エッジルータの網内アドレスを単位とした複数の異なるキューを有している。

次に、タイマー部 305 は次のような理由で設けられている。すなわち、読み出し制御部 303 は、バッファ部 304 上の各キューの中に前回スーパーパケットを送出してから一定時間 T が経過しているキューがあると、スーパーパケット長に達するだけの IP パケットが当該キュー上に溜まっていなくとも、当該キューに関してキュー上に蓄積されている全ての IP パケットを当該キューから読み

出してスーパーパケットを構築し、レイヤ1/レイヤ2処理部307を通じてネットワーク106側に転送する。したがって、この場合に構築されるスーパーパケット110では、スーパーパケット長に満たない部分の空き領域が読み出し制御部303によってPadding されることになる。

【0096】

ここで、上述した一定時間Tは次のようにして決定される。すなわち、実際にネットワークを運用する場合には個々の網内フローに対して最小帯域を設定して転送を行うことが一般的に行われている。そのためエッジルータ装置には、ある固定長のフレームを転送する際の最小帯域が例えば64 kbps, 1.5Mbps, 10Mbpsのように予め設定するようにしている。例えばいま仮に或る網内フローについて最小帯域として10Mbpsを設定するものとし、またスーパーパケットのパケットペイロードの長さが例えば1800byteであるものとし、さらに個々の伝送路の物理速度がOC48 [ITU (国際電気通信連合) の伝送路速度規格であって2.48832Gbps] であるものとする。

【0097】

こうした条件下で10Mbpsの論理速度を満たすためには、一定時間T (秒) $= 1800 \text{ byte} \div 10 \text{ Mbps} = 1800 \times 8 \text{ bit} \div 10 \text{ M} = 14.4 \text{ m秒}$ 毎に一つのスーパーパケットを転送する必要がある。したがって、この場合における上記網内フローの滞留時間の上限値は14.4m秒となり、この上限値を越えていればパケットペイロードに空きが生じていてもスーパーパケットを転送することになる。こうすることで、どんなにトラヒックが混雑していてもユーザに対して10Mbpsの帯域を保証することができる。

【0098】

このほか、上述したような最小帯域の規定がないものの最大遅延時間の規定がある場合には、この最大遅延時間に基づいて一定時間Tをトラヒック (網内フロー) 毎に決定することが考えられる。いま例えば、常に100m秒以内で網内転送を終了させたいトラヒック (網内フロー) が存在しているものとする。ここで、トラヒックの網内遅延は、網内遅延=スーパーパケット構築部205のキュー滞留時間上限値 (送信元エッジルータ側) + 伝送路遅延 + スーパーパケット復元

部 206 のキュー滞留時間上限値（宛先エッジルータ側）で規定される。したがって、こうした網内遅延の最大許容値（この場合は 100 m 秒）をもとに、伝送路遅延および後述するスーパーパケット復元部 206 のキュー滞留時間上限値を加味しつつ、スーパーパケット構築部 205 のキュー滞留時間の上限値を決定して、その値から上述した一定時間 T を決めてゆく。

【0099】

以上のようにすることで、トラヒックが少ない場合であっても、タイムアウト時間に相当する一定時間 T が経過すれば送信元エッジルータから宛先エッジルータに対してスーパーパケットが転送されることになって、網内フロー 107 毎に最小帯域を保証することが可能となる。したがって、タイマー部 305 はバッファ部 304 上のキュー毎にそれぞれ対応した独立のタイマーを内蔵している。読み出し制御部 303 はスーパーパケットを送出した時点でこの旨を宛先エッジルータアドレスとともにタイマー部 305 に通知する。

【0100】

そこでタイマー部 305 はこの通知時点から、通知された宛先エッジルータアドレスのキューに対応するタイマーの計時をゼロから開始させ、上記一定時間 T が経過した時点でタイムアウトしたキューに対応する宛先エッジルータアドレスを読み出し制御部 303 に通知する。ちなみに、上述した処理において、網内フローにトラヒックが全くないことからキューに IP パケットが全く溜まっていない場合、読み出し制御部 303 はスーパーパケットを構築してネットワーク 106 側に送出することはしない。なお、これ以外の場合であっても、タイマー部 305 は読み出し制御部 303 から指示があれば内蔵のタイマーをリセットしてゼロから計時を開始させることができる。

【0101】

次に、ポインタ部 306 は IP パケットが図 5（a）に示したように複数のスーパーパケットに分割される場合に、当該 IP パケットの残部の長さ（同図に示した長さ L_2 ）を保持するための記憶手段である。前述したように、図 5（a）では長さ L_2 の値がスーパーパケット A2 のポインタフィールド①（図 3 参照）として格納されることからポインタ部の名称を冠してある。

次に、レイヤ1／レイヤ2処理部307は、読み出し制御部303から転送されてくるIPパケットに対してレイヤ2ヘッダ34（図17参照）およびレイヤ2トレーラ35を付与してレイヤ2フレーム33に変換したのち、このレイヤ2フレーム33に対してレイヤ1ヘッダ31をさらに付与してレイヤ1フレーム42を構築してネットワーク106側の伝送路100に出力する。

【0102】

（c）スーパーパケット復元部の構成

次に、図11を参照してスーパーパケット復元部206の詳細について説明する。図示した通り、スーパーパケット復元部206はレイヤ1／レイヤ2処理部501、宛先エッジルータ判定部502、書き込み制御部503、バッファ部504、読み出し制御部505、タイマー部506、従来技術で説明したものと同一パケット分割部16（図16を参照）から構成されている。そして、レイヤ1／レイヤ2処理部501は図1に示した伝送路100に接続され、パケット分割部16は図7に示したスイッチ部203に接続されている。

【0103】

まず、レイヤ1／レイヤ2処理部501は、ネットワーク106側（すなわち伝送路100）からスーパーパケット110の含まれたレイヤ1フレームを受信してレイヤ1／レイヤ2の終端処理を施す。すなわち、レイヤ1／レイヤ2処理部501は、レイヤ1フレーム30（図17）のレイヤ1ペイロード32からレイヤ2フレーム33を取り出した後、このレイヤ2フレーム33のレイヤ2ペイロード36部分に載せられているスーパーパケット110を取り出し、これを宛先エッジルータ判定部502に送出する。

【0104】

次に、宛先エッジルータ判定部502は、自身の属するエッジルータ装置に付与されている網内アドレス（自エッジルータアドレス）を内部に保持している。そして宛先エッジルータ判定部502は、レイヤ1／レイヤ2処理部501から転送されてくるスーパーパケット110のパケットヘッダ111に記載されている宛先アドレスフィールドの内容（すなわち、宛先エッジルータアドレス）を監視しており、この宛先エッジルータアドレスが自エッジルータアドレスに一致し

ないスーパーパケットを廃棄する一方、両者が一致している場合には転送されてきたスーパーパケットを書き込み制御部 503 に転送する。

【0105】

次に、書き込み制御部 503 は、図 12 に示したように、バッファ部 504 を利用してネットワーク 106 内の送信元エッジルータの網内アドレスを単位としたキューを管理しており、宛先エッジルータ判定部 502 から転送されてくるスーパーパケット 110 をそのパケットヘッダ 111 中の送信元アドレスフィールドに記載されている送信元エッジルータアドレスに対応したキューへ格納する。

次に、バッファ部 504 は上述したように宛先エッジルータ判定部 502 から書き込み制御部 503 を経由して転送されてくるスーパーパケット 110 を格納するためのメモリであって、このスーパーパケット 110 のパケットヘッダ 111 に含まれる送信元エッジルータアドレスを単位とした複数の異なるキューを有している。

【0106】

次に、読み出し制御部 505 は、上述したバッファ部 504 上の各キューに格納されている各スーパーパケット 110 のパケットペイロード 112 から IP パケットを順次取り出してパケット分割部 16 に送出する。その際、図 5 (a) で示したように IP パケットが複数のスーパーパケットにまたがっている場合は、当該 IP パケットを構成するデータがすべて当該キューに蓄積してから当該 IP パケットを復元してパケット分割部 16 に転送する。

【0107】

例えば図 5 (a) の場合には、スーパーパケット A1 の最後尾にある長さ L1 の領域とスーパーパケット A2 の先頭にある長さ L2 の領域が揃った時点で IP パケット PF を送出するようにする。そのために読み出し制御部 505 は、スーパーパケット A1 のパケットペイロード 112 内のポインタフィールド①及びパケット数フィールド②に基づいて最後に格納されている IP パケットの先頭位置を特定したのち、Padding 有無フィールド③を調べて、当該フィールドの内容が“Padding 無”であってなお且つ当該 IP パケットがスーパーパケット A1 にちょうど収まるものでないときに上述した処理を行うようにする。

【0108】

また、図6に示したように単一のIPパケット37が複数のスーパーパケットにまたがっている場合にはパケットペイロード112の単一パケット占有フィールド④が“ON”になっているので、読み出し制御部505はこうしたスーパーパケットを検出した場合、単一パケット占有フィールド④が“ON”となっているスーパーパケットが連続して到来している間はこれらをバッファ部504上のキューに順次蓄積してゆく。

【0109】

そして読み出し制御部505は、単一パケット占有フィールド④が“OFF”になっているスーパーパケットが到来した時点で、当該スーパーパケットを含めてそれまでにキュー上に蓄積させていたスーパーパケットのパケットペイロード112に詰め込まれているデータを時系列順に繋げて元のIPパケットを復元する。なお、上述したように単一パケット占有フィールド④が“ON”であれば①～③のフィールドはすべて意味を為さなくなるため、読み出し制御部505は①～④のフィールドのうちの単一パケット占有フィールド④を最初に検査するようにして、当該フィールドが“ON”であれば他のフィールドに依らず上述した単一パケットに関わる処理を行うようにしている。

【0110】

次に、タイマー部506は次のような理由から設けられている。すなわち、上述した読み出し制御部505によるIPパケットの復元処理において、一部のスーパーパケットがネットワーク106内で廃棄されてしまうと、いつまで経ってもIPパケットを復元できなくなる場合が起こりうる。そうした場合、正常に転送されてきていたスーパーパケットがバッファ部504内に滞留してしまうことになる。

【0111】

例えば、図5(a)に示したスーパーパケットA2がネットワーク106内で廃棄されてしまうと、スーパーパケットA1上でパケットPFの一部である長さL1の領域に格納されたデータは、図1に示した宛先エッジルータ101g内のバッファ部504に滞留してしまう。こうしたことから、読み出し制御部505

は上記のような状態を監視して、一定時間以上にわたってバッファ部504に滞留しているIPパケット復元待ちのデータがあれば、これを積極的にバッファ部504上のキューから取り出して廃棄している。これによって、バッファ部504上のキューにパケットデータが何時までも滞留するのを防止している。

【0112】

タイマー部506はこの一定時間を計時するためのものであって、バッファ部504上の各キューに対応する独立したタイマーを内蔵している。そしてタイマー部506では、読み出し制御部505がバッファ部504上のあるキューからIPパケットを復元する度にこの旨が当該キューに対応する送信元エッジルータアドレスとともにタイマー部506に通知されると、この通知時点から通知された送信元エッジルータアドレスに対応するキューのタイマーの計時をゼロから開始させる。そして上記一定時間が経過した場合、タイマー部506はタイムアウトしたタイマーに対応するキューの送信元エッジルータアドレスを読み出し制御部505に知らせるようにしている。なお、これ以外の場合であっても、タイマー部506は読み出し制御部505から指示があれば内蔵のタイマーを再度ゼロから計時させるようになっている。

【0113】

ここで、タイマー部506が計時する一定時間は、スーパーパケット構築部205内のタイマー部305に設定した一定時間Tをもとに決定される。というのも、上述したように最小帯域が決定されているのであれば、送信元エッジルータ側でトラヒックが有りさえすればこの最小帯域で決まる時間内にスーパーパケットが1個は必ず到来することが期待できる。したがって、複数のスーパーパケットに分割されたIPパケットが仕掛けり状態となっている場合にも、上記最小帯域で決まる時間だけ待っていればIPパケットの後半部分が届くはずである。もっとも、ネットワーク内で輻輳が生じたりジッタが発生したりするとこの時間が若干前後することになることから、厳密に言えばこうした変動による時間を加味することが好ましい。

【0114】

結局、バッファ部504上に構築されたキューに到着するスーパーパケットの

間隔は、スーパーパケット構築部 205 内で規定される上限時間+網内の遅延分散時間〔例えば CDV (セル遅延変動許容値)〕となる。そして本実施形態では、タイマー部 506 が計時する一定時間の上限値は、このスーパーパケットの到着間隔に対して予め定めておいた保護時間 α を加えたものとしている。つまり、上記一定時間は、「スーパーパケット構築部 205 内で規定される上限時間+網内の遅延分散時間+保護時間 α 」として見積もることができる。もっとも、実際にネットワークを運用するにはこうした時間の見積もりが煩雑である場合も考えられるため、そうした場合は最悪のケースを考慮して十分長くとした時間値を上記一定時間にしてしまっても良い。

【0115】

なお、上述したようにスーパーパケット構築部 205 側の一定時間 T が最小帯域ではなく遅延時間で規定されている場合には、最小帯域から決まる時間の代わりにこの遅延時間の規定値を用いるようにすれば良い。

最後に、パケット分割部 16 は、読み出し制御部 505 から転送されてくる IP パケットを固定長の装置内セルに分割してからスイッチ部 203 に送出し、同時に、転送された IP パケットをアクセスインタフェース部 202-1~202-M の何れに転送すべきかをスイッチ部 203 に通知する。

本実施形態では、以上説明したような仕組みによって、ネットワーク 106 内で固定長のスーパーパケット 110 を運用可能としているほか、スーパーパケット 110 のパケットペイロード 112 に IP パケットを詰め込んで、空き帯域なくリンク利用率 100% で運用することを可能にしている。

【0116】

(4) 動作の説明

次に、上記構成によるネットワーク内の各部の動作について図 13~図 14 に示されるフローチャートを参照しながら詳細に説明する。ここでは、図 1 に示したように、送信元ホスト 105c が発生させた IP パケット 37 の含まれるレイヤ 1 フレームが、網内アドレス“3”を持つ送信元エッジルータ 101c でスーパーパケット 110 の含まれたレイヤ 1 フレームに変換され、これが網内アドレス“10”のコアルータ装置 102b、網内アドレス“9”のコアルータ装置 1

02dに中継されてゆき、網内アドレス“7”を持つ宛先エッジルータ101gにおいて元のIPパケットが含まれたレイヤ1フレームに変換され、宛先ホスト105gに転送されるまでを説明する。

【0117】

まず、ユーザLAN104cに属する送信元ホスト105cがIPパケット37の含まれたレイヤ1フレームを発生させて送出すると、ユーザLAN104cは送出されたIPパケット37をネットワーク106を経由させて転送すべきか否かを判定する。もし、転送すべきでないと判断した場合、ユーザLAN104cは自身に接続された送信元ホスト105c以外の他のホストにIPパケット37が含まれたレイヤ1フレームを送信する。しかしここでは、当該IPパケットがネットワーク106を介して転送されるべきであると判断されたものとする。

【0118】

そのため、ユーザLAN104cはIPパケット37が含まれたレイヤ1フレームをアクセス回線103cを介して送信元エッジルータたるエッジルータ装置101cに転送する。これによって、エッジルータ装置101cに相当する図7のエッジルータ装置201では、アクセスインタフェース部202-1~202-Mのうち、アクセス回線103cに接続されているアクセスインタフェース部に対してIPパケット37の含まれたレイヤ1フレームが入力される（以上、ステップS1）。

【0119】

IPパケット37の含まれたレイヤ1フレームが入力されると、まずは従来技術で説明したのと同様の動作が行われる。すなわち、上記何れかのアクセスインタフェース部は、IPパケット37が含まれたレイヤ1フレームに対してレイヤ1処理、レイヤ2処理を順次実行してIPパケット37を取り出したのち、取り出されたIPパケット37に対してIPレイヤ処理を実行したのち、これを固定長の装置内セルに分割してスイッチ部203に転送する。このとき同時に、アクセスインタフェース部は、IPパケット37のヘッダ部分に含まれている宛先アドレスに従って経路情報を検索して次にIPパケットを転送すべき転送先としてコアルータ装置102bを特定し、スーパーパケット処理部204-1~204-P

のうちコアルータ装 1 0 2 b に接続されている何れかのスーパーパケット処理部を特定してこの旨をスイッチ部 2 0 3 に通知する。

【0 1 2 0】

これにより、スイッチ部 2 0 3 は自身に内蔵されている（図 1 6 と同様の構成をした）クロスバースイッチ部におけるスイッチ接続の適性な組み合わせを計算したのち、アクセスインタフェース部から転送されてくる装置内セルをクロスバースイッチ部に転送してゆく。これによって、装置内セルがクロスバースイッチ部を介して上記特定されたスーパーパケット処理部内のスーパーパケット構築部 2 0 5 に順次入力される。

【0 1 2 1】

すると、エッジルータ装置 1 0 1 c 内のスーパーパケット構築部 2 0 5 では、図 8 に示したパケット復元部 1 8 がスイッチ部 2 0 3 でスイッチングされた装置内セルを積み上げて IP パケット 3 7 を復元して宛先エッジルータ判定部 3 0 1 に送出する。宛先エッジルータ判定部 3 0 1 は、パケット復元部 1 8 から送られた IP パケット 3 7 のヘッダ部分に含まれている宛先 IP アドレスに基づいてアドレス解決テーブル 3 0 8 を検索してアドレス解決を行って、ネットワーク 1 0 6 内の何れの宛先エッジルータに転送すべきかを決定するとともに、当該宛先エッジルータまで IP パケット 3 7 を転送するための経路を判定する。こうして、宛先エッジルータ判定部 3 0 1 はアドレス解決テーブル 3 0 8 からネットワーク 1 0 6 内でスーパーパケット 1 1 0 が最終的に転送される宛先エッジルータ 1 0 1 g の網内アドレス “7” を取得し、この網内アドレスの値を転送されてきた IP パケット 3 7 と一緒にして書き込み制御部 3 0 2 へ転送する。

【0 1 2 2】

書き込み制御部 3 0 2 は、宛先エッジルータ判定部 3 0 1 から転送されてくる IP パケット 3 7 を通知された宛先エッジルータ 1 0 1 g の網内アドレスに対応したバッファ部 3 0 4 上のキューに書き込む（ステップ S 2）。同時に書き込み制御部 3 0 2 は、当該 IP パケット 3 7 のヘッダに記述されているトータル長を参照して IP パケット 3 7 のパケット長を宛先エッジルータアドレスの値 “7” とともに読み出し制御部 3 0 3 に通知する。こうして、パケット復元部 1 8、宛

先エッジルータ判定部 301 及び書き込み制御部 302 は、送信元ホスト 105c から IP パケット 37 の含まれたレイヤ 1 フレームが送出される度に、上述した動作を繰り返して（ステップ S3 およびステップ S4 の判断結果がいずれも “NO”）バッファ部 304 上のキューに IP パケット 37 を蓄積してゆく。

【0123】

一方、読み出し制御部 303 はバッファ部 304 上の各キューに関するキュー長を監視しているため、キュー長がスーパーパケット長に到達したキュー（ステップ S4 の判断結果が “YES”）に関しては当該キュー上の IP パケットをスーパーパケットとして転送するようにする。すなわち、読み出し制御部 303 はキュー長がスーパーパケット長以上になったキューからスーパーパケット長 “n×L” 分の IP パケットを読み出す（ステップ S5 の判断結果が “NO” となった後のステップ S6）。もっとも正確には、上述したようにパケットヘッダ 111 および①～④のフィールドに必要となる長さをスーパーパケット長から減じた長さだけの IP パケットをキューから読み出すことになる。

【0124】

次に、読み出し制御部 303 はバッファ部 304 から読み出した IP パケットをパケットペイロード 112 と見なし、その先頭にパケットヘッダ 111 を付与してスーパーパケット 110 のフレームに変換するとともに、パケットヘッダ 111 上の宛先アドレスフィールドおよび送信元アドレスフィールドへそれぞれ宛先エッジルータアドレス（すなわち “7”）および送信元エッジルータアドレス（すなわち “3”）を格納する。また、読み出し制御部 303 はパケットペイロード 112 上のポインタフィールド①に対して、パケットペイロード 112 に現れる最初の IP パケットの先頭位置を示すポインタ値を書き込む。さらに、読み出し制御部 303 はパケットペイロード 112 のパケット数フィールド②に対して、パケットペイロード 112 上に存在する IP パケットの先頭位置の個数を書き込む。

【0125】

このほか、読み出し制御部 303 はパケットペイロード 112 の最後の IP パケットが複数のスーパーパケットにまたがっているか否か（ステップ S5 の判断

結果がそれぞれ“YES”または“NO”)に応じて Padding有無フィールド③を「Padding有」又は「Padding無」にそれぞれ設定する。また、読み出し制御部 303 はいま構築したスーパーパケット 110 のパケットペイロード 112 全体が単一の IP パケット 37 のデータで占有されている場合 (図 6 の場合) には、パケットペイロード 112 の単一パケット占有フィールド④を“ON”に設定し、さもなくば当該フィールドを“OFF”に設定する。そして、読み出し制御部 303 は構築したスーパーパケット 110 をレイヤ 1 / レイヤ 2 処理部 307 に転送する。

【0126】

ここで、スーパーパケット長は固定長であるのに対して IP パケットは可変長であることから、読み出し制御部 303 がキューからスーパーパケット長 (正確には上述した長さ) 分の IP パケットを取り出した場合に最後に挿入される IP パケットが部分的に当該キュー上に残される可能性がある [図 5 (a) のパケット PA ないし PF]。なお、この場合読み出し制御部 303 は当該 IP パケットを分割して例えば図 5 (a) に示すパケット PF の前半部分までの IP パケットのデータからスーパーパケット A1 を構築し、パケット PF の残部についてはその後に行われるスーパーパケット A2 の構築の際に使用するためにこれをキュー上に残すようにする (ステップ S14)。そこでこうした場合、読み出し制御部 303 はキュー上に残される残部の長さを算出してこの値をポインタ部 306 に保持させる。なお、残部が存在しない場合、読み出し制御部 303 はポインタ部 306 に“0”を格納する。これによって、ポインタ部 306 に保持されている残部の長さ [図 5 (a) で言えば長さ L2] が、次に構築されるスーパーパケット [図 5 (a) のスーパーパケット A2 に相当] のポインタフィールド④の値として使用されることになる。

【0127】

このほか、トラヒックが少ないためにキュー長がなかなかスーパーパケット長 (正確には上述した長さ) に達しないキューに関しては、一定時間 T 以上の時間が経過してもスーパーパケットが読み出し制御部 303 によって構築されない。このため、読み出し制御部 303 が前回スーパーパケットを構築してから計時を

開始していたタイマー部 305 内のタイマーは一定時間 T の経過によってタイムアウトする（ステップ S3 の判断結果が “YES”）。その結果、このタイムアウトしたキューに対応する宛先エッジルータアドレスがタイマー部 305 から読み出し制御部 303 に通知される。

【0128】

すると読み出し制御部 303 は、当該宛先エッジルータアドレスで指定されるキューのキュー長がスーパーパケット長（正確には上述した長さ）に達していなくても当該キューに蓄積されている IP パケットから上記同様にスーパーパケット 110 を強制的に構築する（ステップ S16）。なお、この場合にはスーパーパケット 110 に必ず空き領域が存在するため、読み出し制御部 303 はパケットペイロード 112 の Padding 有無フィールド③を「Padding 有」に設定する。このほか、読み出し制御部 303 はタイマー部 305 に指示を行ってタイムアウトしたキューに対応するタイマーを再度ゼロから計時させるようにする。

【0129】

次に、レイヤ 1/レイヤ 2 処理部 307 は読み出し制御部 303 からスーパーパケット 110 が転送されてくると、これに必要なレイヤ 1 ヘッダ 31, レイヤ 2 ヘッダ 34, レイヤ 2 トレーラ 35 を付与してレイヤ 1 フレームに変換し、これを伝送路 100 へ転送する（ステップ S7）。こうして、送信元エッジルータ 101c からネットワーク 106 に対してスーパーパケット 110 の含まれたレイヤ 1 フレームが送出され、伝送路 100 を介してコアルータ装置 102b に転送される。するとコアルータ装置 102b はエッジルータ装置 101c から転送されてきたレイヤ 1 フレームに含まれているスーパーパケット 110 のパケットヘッダ 111 に記述された宛先アドレスフィールドの宛先エッジルータアドレスに従ってコアルータ装置 102d に転送する。

【0130】

ここで、コアルータ装置は従来技術で説明したルータ装置と同様の構成であるほか、コアルータ装置にとってはレイヤ 1 フレームに含まれているスーパーパケット 110 が通常の IP パケット 37 と同じように見える。したがって、コアルータ装置 102b は従来技術で説明したのと同様にして、スーパーパケット 11

0が含まれたレイヤ1フレームをコアルータ装置102dに転送する。もっとも、従来のルータ装置では装置内セル単位ないし固定長L毎にスイッチ動作が発生していたのに対し、本実施形態のルータ装置ではスーパーパケット単位ないしスーパーパケット長(=長さ“ $n \times L$ ”)毎にしかスイッチ動作が発生しない。このため本実施形態ではスイッチング回数を従来の $1/n$ に低減させることができる。そして、コアルータ装置102dもコアルータ装置102bと全く同様の手順に従って、転送されてくるレイヤ1フレームをエッジルータ装置101gに転送する。このようにして、スーパーパケット110に含まれたレイヤ1フレームは以上説明した中継を経て宛先エッジルータ101gに到達する(以上、ステップS8)。

【0131】

次に、宛先エッジルータ101gでは、スーパーパケット110に含まれたレイヤ1フレームが伝送路100を通じてスーパーパケット処理部204-1~204-Pの何れかに具備されたスーパーパケット復元部206に入力される。スーパーパケット復元部206では、まずレイヤ1/レイヤ2処理部501がレイヤ1フレームを終端し、当該終端処理によって得られたスーパーパケット110を宛先エッジルータ判定部502に転送する。宛先エッジルータ判定部502は、転送されたスーパーパケット110のパケットヘッダ111中の宛先アドレスフィールドに記述された宛先エッジルータアドレスを参照して自エッジルータアドレス(すなわち“7”)と比較する。その結果、もし両者が一致していなければ、転送されてきたスーパーパケット110は誤ってルーティングされてきたものであるので、宛先エッジルータ判定部502は転送されたスーパーパケット110を廃棄する。もっともこの場合は両者が一致しているため、宛先エッジルータ判定部502は転送されてきたスーパーパケット110を書き込み制御部503へそのまま転送する。

【0132】

書き込み制御部503は、転送されてきたスーパーパケット110のパケットヘッダ111を参照して、その送信元アドレスフィールドに記載されている送信元エッジルータアドレスに対応するバッファ504上のキューへ転送されてきた

スーパーパケット 110 を書き込む（ステップ S9）。同時に書き込み制御部 503 は、当該スーパーパケット 110 のパケットペイロード 112 に含まれている①～④のフィールドを参照して、バッファ部 504 上のキューに書き込んだスーパーパケット 110 に含まれる IP パケットの数、図 5（a）のスーパーパケット A1 に含まれるパケット PF に相当する「IP パケット復元処理の仕掛かり状態」にあるパケット（以下、「仕掛かり状態パケット」という）の有無等に関する情報を作成して、これらの情報を読み出し制御部 505 に通知する。

【0133】

読み出し制御部 505 は、書き込み制御部 503 から通知された情報に基づき、バッファ部 504 上の各キューに格納されているスーパーパケット 110 から IP パケットを順次取り出して、これら IP パケットをパケット分割部 16 に転送してゆく（ステップ S10 及びステップ S11 の判断結果がいずれも“NO”となった後のステップ S12）。その際、読み出し制御部 505 はパケットペイロード 112 上のポインタフィールド①、パケット数フィールド②および各 IP パケットのヘッダに記述されているトータル長に従って各々の IP パケットの先頭位置を特定して IP パケットを取り出してゆく。そして、パケット分割部 16 は読み出し制御部 505 から転送されてくる IP パケットを装置内セルに分割してスイッチ部 203 に送出する。なお、送信元エッジルータ 101c がスーパーパケットを構築する際にタイムアウトが発生して強制的にスーパーパケットを構築したような場合には、スーパーパケット 110 のパケットペイロード 112 中の Padding 有無フィールド③が「Padding 有」となるが、この場合も上記と同様である。

【0134】

ここで、スーパーパケットから IP パケットを取り出してゆく過程で仕掛かり状態パケットを検出する（ステップ S10 の判断結果が“NO”、かつ、ステップ S11 の判断結果が“YES”）と、読み出し制御部 505 はこの仕掛かり状態パケットをバッファ部 504 上のキューに格納したまますぐには読み出しを行わず、現在処理中のスーパーパケット〔図 5（a）で言えばスーパーパケット A1〕に続くスーパーパケット〔図 5（a）で言えばスーパーパケット A2〕が到

着してからIPパケットを復元するようにする（ステップS15）。すなわち、読み出し制御部505はスーパーパケットA1が転送されてきた時点で、先行するスーパーパケットとの間にまたがっているパケットPAを除き、パケットPB～PEを順次取り出してパケット分割部16に送出する。

【0135】

一方、パケットPFにつき、読み出し制御部505は長さL1の領域のデータをバッファ部504のキュー上に保持しておき、スーパーパケットA2が転送されてきた時点で、パケットペイロード112のポインタフィールド①の直後から当該フィールドで示される位置の直前の位置までのデータ（即ち、長さL2の領域）を取り出し、これをバッファ部504上に保持してある長さL1のデータの後ろに繋げて元のパケットPFを復元したのち、このパケットPFをパケット分割部16に転送する。

【0136】

またこの場合において、スーパーパケットA2が網内フロー107の途中で廃棄されてしまうと、読み出し制御部505によって仕掛かり状態パケットが復元されないままとなる。しかしこのときタイマー部506は、読み出し制御部505からIPパケットが取り出される度に、当該IPパケットの属するキューに対応した内部のタイマーの計時をやり直している。このためタイマー部506は、当該タイマーがタイムアウトした時点で、タイムアウトの旨を当該タイマーに対応する送信元エッジルータアドレスとともに読み出し制御部505に通知する（ステップS10の判断結果が“YES”）。

【0137】

この結果、読み出し制御部505は送信元エッジルータアドレスに対応するバッファ504上のキューに残留している仕掛かり状態パケットを廃棄する（ステップS17）。またこの後に読み出し制御部505がタイムアウトしたキューに蓄積された新たなスーパーパケットからIPパケットを取り出すようになると、タイマー部506は当該キュー対応するタイマーを再度ゼロから計時させるようにする。以上のようにして、本実施形態ではバッファ部504上のキューにごみパケットが残留するのを防いでいる。

【0138】

そして、スーパーパケット復元部206の読み出し制御部505から転送されたIPパケットは、パケット分割部16で装置内セルに分割されたのち、図7に示したスイッチ部203に入力され、以後は、従来技術で説明したのと同様の動作が行われる。すなわち、パケット分割部16は読み出し制御部505から転送されてくるスーパーパケットを装置内セルに分割してスイッチ部203に送出し、これと同時に、当該スーパーパケットをアクセスインタフェース部202-1～202-Mの何れに転送するかを決定してスイッチ部203に通知する。そして、この場合は次に転送すべき転送先としてユーザLAN104gが特定され、アクセスインタフェース部202-1～202-MのうちでユーザLAN104gに接続されているアクセスインタフェース部が特定される。

【0139】

そこで、スイッチ部203は転送された装置内セル401をアクセスインタフェース部毎に設けられた内蔵メモリ（図示省略）に格納してゆくとともに、内部のクロスバースイッチ部におけるスイッチ接続の適性な組み合わせを計算して、スーパーパケットを構成する装置内セル40をこのクロスバースイッチ部に転送してゆく。これにより、クロスバースイッチ部を介して装置内セル40が、アクセスインタフェース部202-1～202-Mのうち、アクセス回線103gに接続されたアクセスインタフェース部に順次転送されてゆく。

【0140】

当該アクセスインタフェース部では、転送されてくる装置内セル40をスーパーパケット処理部毎に異なる内蔵メモリへ格納してゆき、装置内セル40を積み上げて元のIPパケットを復元する。この後、当該アクセスインタフェース部は復元されたIPパケットに対してIPレイヤ処理、レイヤ2処理、レイヤ1処理を順次行って当該IPパケットの含まれたレイヤ1フレームを構築してアクセス回線103gを介してユーザLAN104gに転送する。すると、ユーザLAN104gはエッジルータ装置103gから転送されてきたレイヤ1フレームに含まれるIPパケットの転送先を判定し、その結果として当該レイヤ1フレームを宛先ホスト105gに転送する（以上、ステップS13）。

【0141】

以上のようにして、送信元エッジルータ101cは、送信元ホスト105cが発生させたIPパケットを含むレイヤ1フレームをユーザLAN104c経由で受け取り、これをそのままネットワーク106に転送するのではなく、ネットワーク106内でのみ定義されるスーパーパケット110に変換してからレイヤ1フレームの形で網内フロー107上を伝送させている。このスーパーパケット110は、通常のIPパケットがネットワーク内を転送されるのと全く同様にしてネットワーク106内に配置されているコアルータ装置を経由して宛先エッジルータ101gまで中継されてゆく。そして、宛先エッジルータ101gがネットワーク106内を転送されてきたスーパーパケット110を元のIPパケットに復元してユーザLAN104gから宛先ホスト105gまで転送している。

【0142】

以上詳述したように、本実施形態では、各ユーザLANから転送されてくるIPパケットを従来の装置内セルの n 倍（ n ：2以上の自然数）の長さを持った固定長のスーパーパケットに詰め込んだ後にネットワークへ転送している。このため、ネットワークを構成しているエッジルータ装置又はコアルータ装置のスイッチング単位を装置内セルの長さからスーパーパケット長にまで大きくすることが可能となり、上述した従来の第1の問題点が解決される。

【0143】

すなわち、従来は、クロスバースイッチ部のスイッチ単位である装置内セルの長さが短いために、クロスバースイッチ部でスイッチ接続を行うためのスケジューラの計算時間も短くなってしまい、ルータ装置のスイッチ規模を大きくすることができなかった。これに対し、本実施形態では装置内セルの大きさそのものを実質的に増やしてスーパーパケット長とすることができる。このため、入力インタフェースや出力インタフェースの数が増えても、スイッチ接続の計算を行う時間（転送時間単位）を十分に確保することができ、結果としてルータ装置のスイッチ容量を増大させることが可能である。また、転送時間単位を増やすことができることから、リンク速度が増加した場合であってもスイッチ接続の計算時間を確保できるため、ルータ装置内のスイッチの拡張が可能となる。

【0144】

また、本実施形態では、ネットワーク上で転送されるパケットの大きさをスーパーパケット長にまで拡大しているため、従来の第2の問題点も解決することができる。すなわち、この従来の問題点は、ネットワーク上で転送されるIPパケットの最小長がルータ装置内の転送単位である装置内セルよりも短いために発生していた。そこで本実施形態では、ネットワーク上で転送されるパケットの最小の大きさをスーパーパケット長としてIPパケットの最小長より十分大きくとるようにして、上述した問題点が発生しないようにしている。

【0145】

さらに、本実施形態では、ネットワーク上で転送されるスーパーパケットの長さを装置内セルの整数倍にしているため、従来の第3の問題点も解決することができる。すなわち、この従来の問題点は、ネットワーク上で転送される可変長のIPパケットの長さが固定長の装置内セルの長さの整数倍になるとは限らないために、IPパケットを装置内セル単位に分割したときに最後に分割される装置内セルにおいて空き帯域が生じることが原因となっている。しかるに、本実施形態ではスーパーパケットを装置内セルに分割しても、分割された装置内セルに空き帯域が発生することは全くないので、従来のような問題が発生することはない。

こうして、本実施形態によれば大規模なルータ装置の構築が可能となり、なおかつ、ネットワーク上で転送されるIPパケットの数を大幅に増大させることができるようになる。

【0146】

なお、上述した説明では、エッジルータ装置101cが送信元エッジルータとして機能し、エッジルータ装置101gが宛先エッジルータとして機能するものとしたが、各エッジルータ装置が送信元エッジルータとしての機能と宛先エッジルータとしての機能を併有していることはもちろんである。

また、上述したように本実施形態ではIPパケットに特化した形で説明を行ったが、IPパケットの代わりにATMセル、STMフレームなどを用いることもできる。すなわち、これらは何れもヘッダ部分にアドレス情報を有するなどIPパケットと同様のフォーマットとなっていることから、上述した説明におけるI

P パケットを ATM セル、STM フレームなどと読み替えれば、以下に述べる点を除いて上述した IP パケットの場合と同様でパケット通信システムを実現することが可能である。

【0147】

すなわち、トラヒックが ATM あるいは STM である場合の相違点として、例えば図 3 に示したスーパーパケット 110 に詰め込まれるパケット PA～PF が何れも固定長になる点が挙げられる。また、ATM セルや STM フレームはいずれも固定長であるため、これらセルやフレームはその内部にセル長ないしフレーム長のフィールドを具備してはいない。そこで、パケットペイロード 112 上で各セルないし各フレームの先頭位置を求めるにあたっては、エッジルータ装置に予め ATM セルや STM フレームの長さに関する情報を定義しておき、エッジルータ装置がこれらの情報をもとにしてそれぞれの ATM セルないし STM フレームの先頭位置を認識するようにする。

【0148】

このほか、IP パケット、ATM セル、STM フレームなど、複数種類のセル、パケット、フレームを混在して転送させる場合は次のようにすれば良い。すなわち、Egress ルータはネットワーク 106 から転送されるスーパーパケットに含まれるのが IP パケット/ATM セル/STM フレームの何れであるのかを認識した上でこれらをユーザ LAN に戻してやれば良い。そのために、各々の Egress ルータに接続された複数のアクセス網のトラヒック種別 (IP/ATM/STM) が異なっているものがある場合、もしくは、同じトラヒックであってもキュー滞留時間や優先制御などの関係で Egress ルータが異なる扱いをする必要があるのであれば、これら異なるトラヒックに対してそれぞれ違う網内アドレスを付与するようにする。そうすることで、例えば Egress ルータは、スーパーパケット 110 のパケットヘッダ 111 に記述されている宛先アドレスの網内アドレス値をもとに上記 3 種類のうちの何れであるのかなどを判断してそれに応じた個別の処理を行うことが可能となる。

【0149】

以上のほか、例えばトラヒックが STM である場合には、図 1 に示した網内フ

ロー 107 は CBR (Continuous Bit Rate) となり、また、図 3 に示したポインタフィールド①は、スーパーパケット 110 に詰め込まれる STM フレームの先頭を指し示すように構成すれば良い。

また、上述した説明ではエッジルータ装置内にスイッチ部を設けた形態を示したが、これに限らず例えば、パケットやフレームをスイッチングする機能ブロックとスーパーパケットの構築・復元を行う機能ブロックをそれぞれ独立した装置で実現するようにしても良い。

【0150】

【発明の効果】

以上説明したように、請求項 1, 12, 17 記載の発明では、ネットワーク上でパケットを中継する際のスイッチ単位となる固定長セルの複数倍の大きさを持った固定長パケットをスーパーパケットとして定義して、送信元のアクセス網が発生させるパケットをスーパーパケットに変換してネットワークに送出するほか、ネットワークからのスーパーパケットに載せられているパケットを取り出して送信元が発生させた元のパケットを復元して宛先のアクセス網へ送出している。これにより、ネットワーク上で転送されるパケットの長さが従来は一般に可変長であったのに対し、上記各請求項記載の発明ではスーパーパケット長に相当する固定値になっていることが保証される。このため、ネットワーク上の中継地点等に配置された中継手段のスイッチング単位を従来の装置内セルの大きさから実質的にスーパーパケットの大きさまで拡張することが可能になる。したがって、中継手段の入出力インタフェース数が増えても、スイッチ接続設定の計算のための時間を十分確保することができるようになってスイッチ容量を増大させることができる。また、中継手段内部における転送時間を増やすことができるため、リンク速度が増加した場合であってもスイッチ接続設定のための計算時間を十分確保することができ、スイッチの拡張が可能となる。また、スーパーパケットの長さを大きくすればそれだけ中継手段におけるスイッチ回数が減少するため、それだけ中継手段を大規模化することが可能である。

【0151】

また、請求項 2, 18 記載の発明では、スーパーパケットを構築するにあたっ

て、宛先側のパケット転送装置毎に用意されたキュー等へパケットを蓄積してゆき、スーパーパケット構築に必要な分だけパケットが溜まった時点で、当該キューからパケットを抽出してスーパーパケットを構築している。このため、宛先別に転送すべきパケットに関するスーパーパケットの構築処理を宛先毎に管理されたキュー上で効率的に行うことが可能となる。

また、請求項 3，19 記載の発明では、スーパーパケットが構築される度に計時を開始する第 1 の計時手段が第 1 のタイムアウト時間の経過を検出したことで、当該計時手段に対応するキューに蓄積されているパケットがあればこれらのパケットをもとにスーパーパケットを構築している。これにより、トラヒックが少ないことが原因となつてなかなかスーパーパケットが構築されないような状況下にあつても、上記第 1 のタイムアウト時間で必ずスーパーパケットがネットワークに送出され、スーパーパケットが構築されずに送信元のパケット転送装置の内部にいつまでも滞留するようなことが無くなる。

【0152】

また、請求項 4 記載の発明では、ネットワーク上のトラヒック毎の最小帯域から第 1 のタイムアウト時間を決定しているため、トラヒックが少ない状況であっても最小帯域を必ず保証することができる。

また、請求項 5 記載の発明では、ネットワーク上のトラヒック毎の許容網内遅延時間から第 1 のタイムアウト時間を決定しているため、当該第 1 のタイムアウト時間内に網内転送が終了することを保証できるようになる。

また、請求項 6，18 記載の発明では、パケットを復元するにあたって、送信元側のパケット転送装置毎に用意されたキュー等へスーパーパケットを蓄積してゆき、当該キューに蓄積されたスーパーパケットから送信元のアクセス網が発生させたパケットを復元している。このため、宛先側のパケット転送装置では送信元のパケット転送装置側からそれぞれ個別に送られてくるパケットを送信元毎に管理することが可能となる。したがって、例えば複数のスーパーパケットに分割されて送られてくるパケットなども効率的に復元することが可能となる。

【0153】

また、請求項 7，20 記載の発明では、パケットが復元される度に計時を開始

する第2の計時手段が第2のタイムアウト時間の経過を検出したことで、当該計時手段に対応するキューに滞留しているスーパーパケットがあればこれを廃棄している。これにより、パケットが複数のスーパーパケットに分割されてネットワーク上を転送されてくる場合などにおいて、一部のスーパーパケットがネットワーク上で廃棄されたことが原因となって、宛先側のパケット転送装置側でいつまで経ってもパケットが復元できずに既に転送済みのスーパーパケットが滞留してしまうことがなくなり、ごみパケットがキュー上に残留することを防止できる。

また、請求項8記載の発明では、ネットワーク上のトラヒックの最小帯域または網内遅延の最大許容値、ネットワーク内の遅延分散時間、既定の保護時間をもとに上記第2のタイムアウト時間を決定している。このため、トラヒックの最小帯域ないし最大網内遅延時間を保証できるとともに、ネットワーク内で輻輳やジッタが発生した場合であってもこれら最小帯域ないし最大網内遅延時間を保証することが可能となる。

【0154】

また、請求項9，21記載の発明では、送信元のパケット転送装置側でスーパーパケットを構築する際に、パケットが複数のスーパーパケットにまたがることを検出して当該パケットをこれら複数のスーパーパケットに分割して格納するようにしている。また、宛先のパケット転送装置側でパケットを復元する際には、スーパーパケット上のパケットが複数のスーパーパケットにまたがっていることを検出して、これら複数のスーパーパケットに分割して格納されているパケットのデータを繋げて送信元のアクセス網が生成した元のパケットを復元している。これにより、どのような長さを持ったパケットをスーパーパケットへ詰め込んでいっても、スーパーパケット上に空き領域が生じることはなくなる。したがって、空き帯域なくリンク率100%でネットワークを運用しつつ、宛先のパケット転送装置側で元のパケットを正しく復元することができる。

【0155】

また、請求項10記載の発明では、スーパーパケットのペイロードが単一のパケットのデータで占有されているか否かを示す単一パケット占有情報をスーパーパケット上に設けている。そして、送信元のパケット転送装置側でスーパーパケ

ットを構築する際は、各パケット長とスーパーパケットのペイロード長との比較から単一パケット占有情報を設定する。一方、宛先のパケット転送装置側では、単一パケット占有情報が占有有であることを検出した場合に、当該情報が占有無のスーパーパケットが到来するまでのこれら一連のスーパーパケット上に載せられているパケットのデータを繋げてゆき、送信元のアクセス網が生成した元のパケットを復元している。これにより、単一のパケットがスーパーパケット長に比べて長大であって、当該パケットが多数のスーパーパケットに分割されて転送されるような場合であっても、空き帯域を生じることなくネットワーク上を転送させて宛先のパケット転送装置側で元のパケットを正しく復元することができる。

【0156】

また、請求項11記載の発明では、スーパーパケットのヘッダをIPパケットなどの従来から用いられているパケットと同一のフォーマットにしている。これにより、ネットワーク上に配置された中継手段を従来から用いられている通常のルータと同様の構成で実現することが可能となる。つまり、中継手段については従来からのルータを流用することができるため、アクセス網とネットワークをつなぐ部分にだけ上述したパケット転送装置を新たに開発して設置すれば、スーパーパケットで運用されるパケット通信システムを構築することが可能となる。

また、請求項13、22記載の発明では、パケット転送装置及び中継手段に対してネットワーク内でのみ定義される固有の網内アドレスを付与しておき、スーパーパケットのヘッダの送信元アドレス及び宛先アドレスには、送信元及び宛先のパケット転送装置に付与された網内アドレスを格納するようにしている。このため、ネットワーク内では個々のパケットのヘッダ部分に含まれているIPアドレスなどの宛先アドレスなどを参照することなく網内アドレスのみで運用することが可能となる。

【0157】

また、請求項14記載の発明では、宛先パケット転送装置が同一のスーパーパケットについて、宛先のアクセス網の種別に応じてスーパーパケットのヘッダ中の宛先アドレスに異なる網内アドレスを割り当てている。これにより、個々のパケット転送装置に接続されるアクセス網として、既存の一般的なアクセス網、A

TM網、STM網などの多様なネットワークを混在させることが可能となる。また、同じトラフィックないしアクセス網の種別であってもキュー滞留時間や優先制御などの関係で異なる扱いをする必要がある場合にも、例えば宛先パケット転送装置がスーパーパケット中の宛先アドレスに応じて個別の処理を行うことが可能となる。

また、請求項15記載の発明では、ネットワーク内の中継手段が、パケット転送装置および中継手段の数に相当するだけのエントリ数を持った経路検索テーブルを有しており、スーパーパケットのヘッダ中の宛先アドレスをもとに当該テーブルを検索して中継先を特定している。これにより、従来のIPルータ装置があらゆる宛先IPアドレスに対応した経路検索テーブルを必要としていたのとは比べた場合、ネットワーク内で定義される少数の網内アドレスに対応したエントリ数の経路検索テーブルを容易すれば良くなり、従来に比べて小規模な構成とすることが可能である。

また、請求項16記載の発明では、送信元パケット転送装置から宛先パケット転送装置までのトラフィックである網内フローが同一であればスーパーパケットのヘッダを同じものとしている。これにより、ネットワーク内のトラフィックの流れを網内フロー別に転送される一連のスーパーパケットを一意に識別することが可能となる。

【図面の簡単な説明】

【図1】 本発明の一実施形態によるパケット通信システムの構成を示すブロック図である。

【図2】 同実施形態において、エッジルータ装置とアクセス網の接続関係の各種形態を例示したブロック図である。

【図3】 同実施形態によるスーパーパケットのフレームフォーマットを示す説明図である。

【図4】 同実施形態によるスーパーパケットを構成するパケットヘッダのフォーマットを示す説明図である。

【図5】 同実施形態においてスーパーパケットにIPパケットが詰め込まれた様子を示す説明図であって、(a)はあるIPパケットが複数のスーパーパ

ケットに分割される場合を示し、(b)はスーパーケットに空き領域が存在する場合を示している。

【図6】 同実施形態において単一のIPケットが複数のスーパーケットにまたがって転送される様子を示した説明図である。

【図7】 同実施形態によるエッジルータ装置の構成を示すブロック図である。

【図8】 図7に示したエッジルータ装置におけるスーパーケット構築部205の構成を示すブロック図である。

【図9】 図8に示したアドレス解決テーブル308のテーブルフォーマット例を示す説明図である。

【図10】 図8に示したバッファ部304上で管理される宛先エッジルータアドレス単位のキューを示す説明図である。

【図11】 図7に示したエッジルータ装置におけるスーパーケット復元部206の構成を示すブロック図である。

【図12】 図11に示したバッファ部504上で管理される送信元エッジルータアドレス単位のキューを示す説明図である。

【図13】 同実施形態によるパケット通信システム内の概略動作を示した第1のフローチャートである。

【図14】 同実施形態によるパケット通信システム内の概略動作を示した第2のフローチャートである。

【図15】 従来技術によるパケット通信システムの構成を示すブロック図である。

【図16】 図15に示した各ルータ装置の構成を示すブロック図である。

【図17】 従来技術において、伝送路又はアクセス回線上で伝送されるレイヤ1フレームがルータ装置内で分解・組立される様子を示した説明図である。

【図18】 従来技術においてIPケットが装置内セルに分割される様子を示した説明図である。

【図19】 従来技術においてIPケットを装置内セルに分割することで最後尾にある装置内セルに空き領域が生じる様子を示した説明図である。

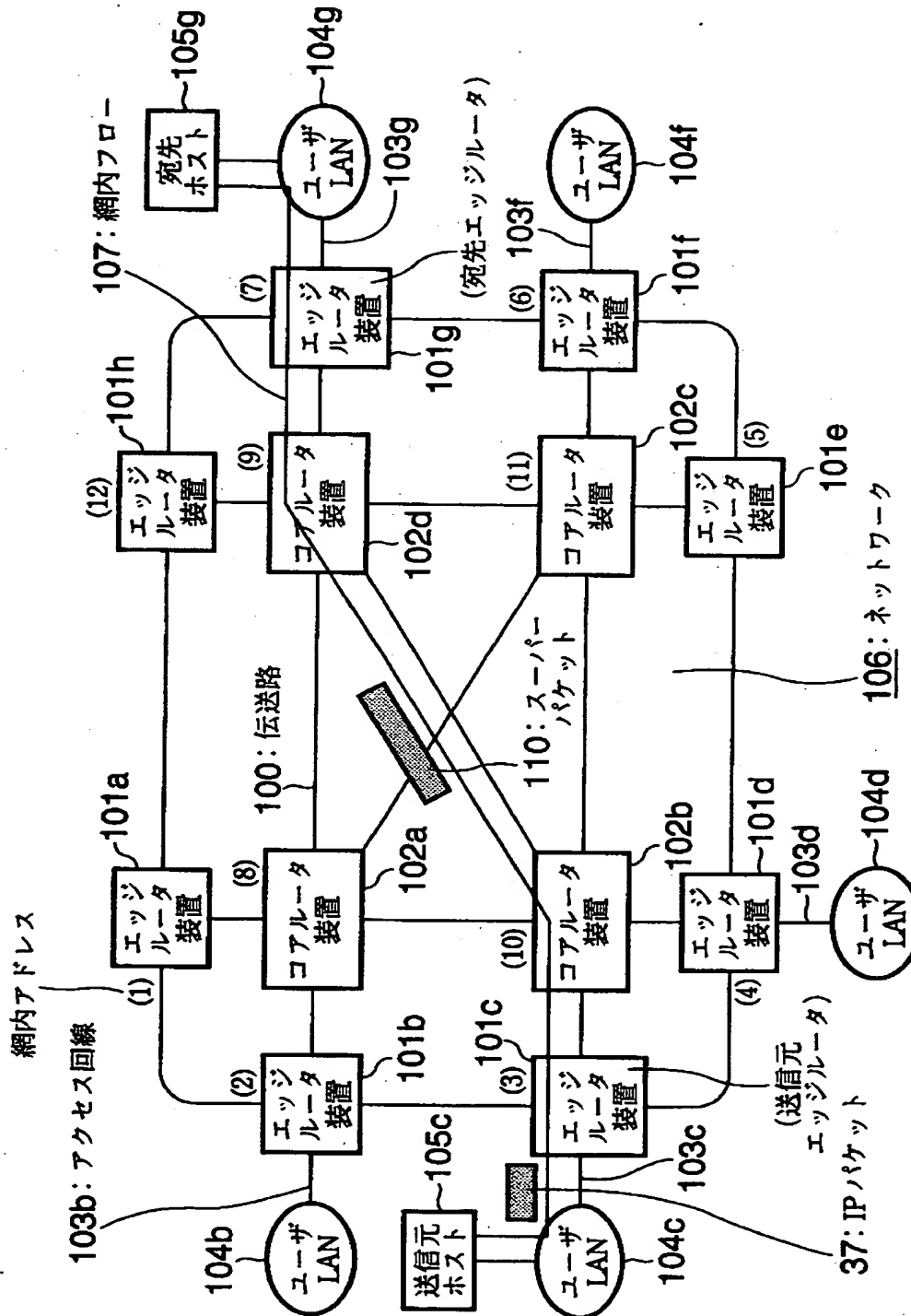
【符号の説明】

16…パケット分割部、18…パケット復元部、37…IPパケット、100…伝送路、101a～101h、101…エッジルータ装置、102a～102d…コアルータ装置、103b～103d、103f、103g…アクセス回線、104b～104d、104f、104g…ユーザLAN、105c…送信元ホスト、105g…宛先ホスト、106…ネットワーク、110…スーパーパケット、111…パケットヘッダ、112…パケットペイロード、202-1～202-M…アクセスインタフェース部、203…スイッチ部、204-1～204-P…スーパーパケット処理部、205…スーパーパケット構築部、206…スーパーパケット復元部、301…宛先エッジルータ判定部、302…書き込み制御部、303…読み出し制御部、304…バッファ部、305…タイマー部、306…ポインタ部、307…レイヤ1／レイヤ2処理部、308…アドレス解決テーブル、501…レイヤ1／レイヤ2処理部、502…宛先エッジルータ判定部、503…書き込み制御部、504…バッファ部、505…読み出し制御部、506…タイマー部

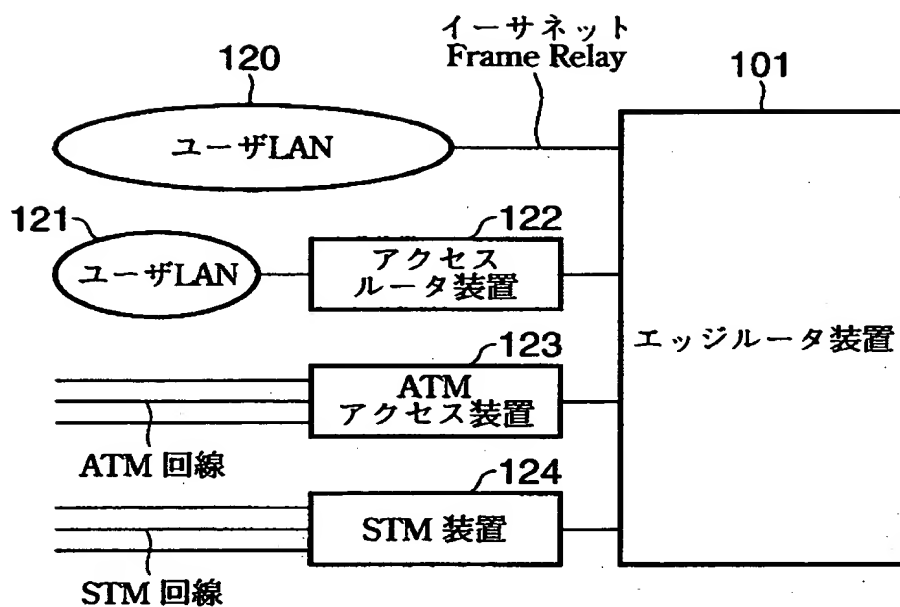
【書類名】

図面

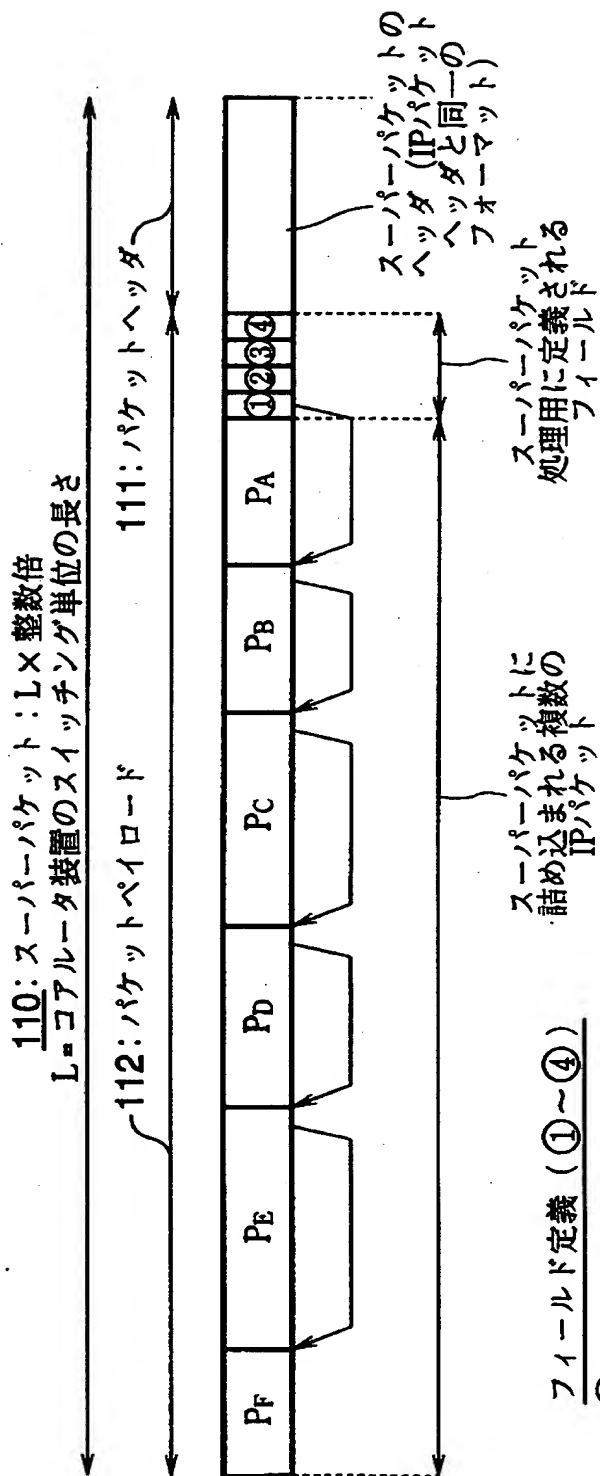
【図 1】



【図 2】



【図 3】



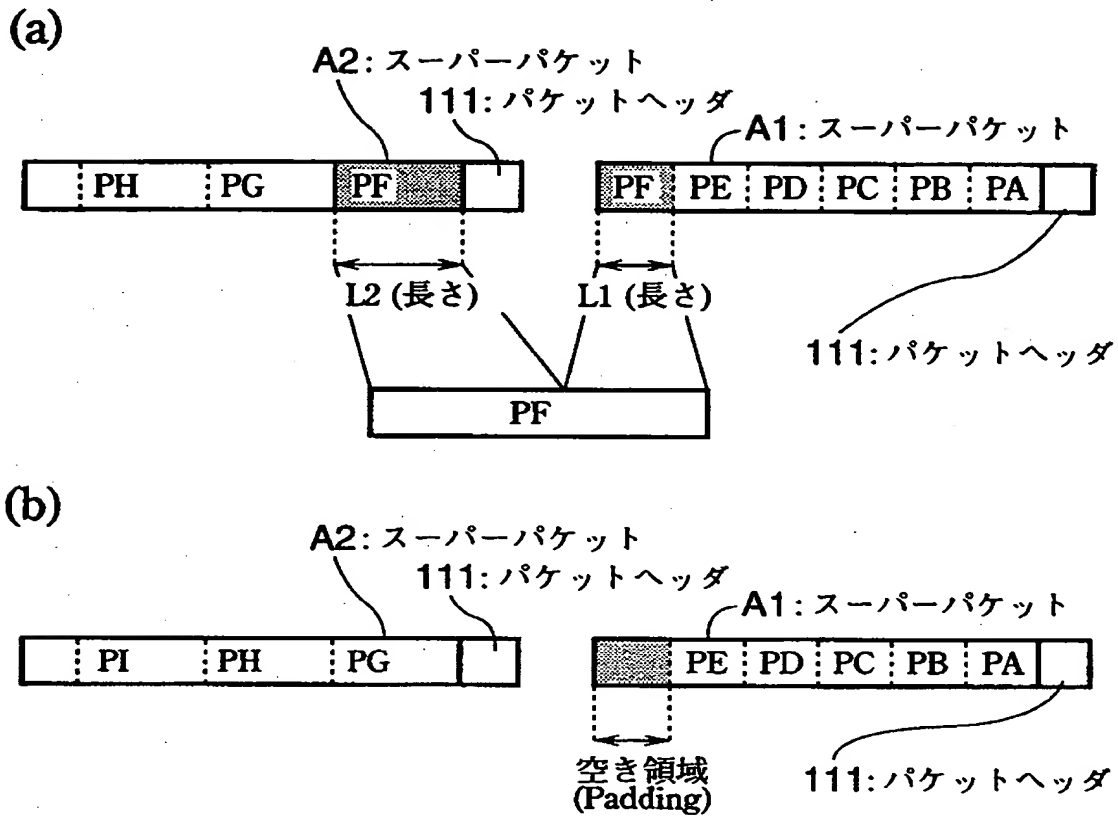
- ① ポインタ フィールド
 スーパーパケットに詰め込まれる最初のIPパケットの先頭位置を示すポインタ
- ② パケット数フィールド
 スーパーパケットに発生するパケット先頭位置の数を示す
- ③ Padding 有無フィールド
 スーパーパケットの最後に現れるIPパケット (PF) が実際のIPパケットのデータか、単なる空き領域にPaddingされているかを示す
- ④ 単一パケット占有フィールド
 本スーパーパケットが単一のIPパケットのデータの一部分ですべて使用され、パケット先頭位置やPaddingがないことを示す

【図 4】

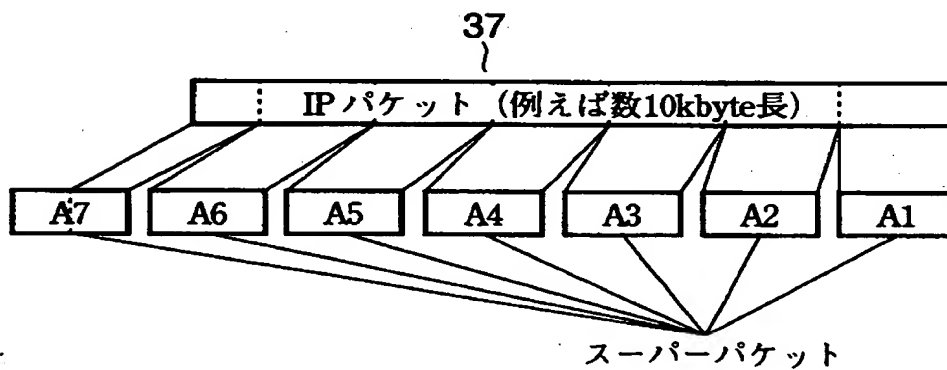
スーパーパケットのヘッダ例 (IPv4パケットを使用した場合)

バージョン	IHL	サービスタイプ	トータル長	
識別子 (ID)		フラグ	フラグメント・オフセット	
生存時間 (TTL)	プロトコル		ヘッダ・チェックサム	
送信元アドレス (送信元エッジルーターアドレスを記入)				
宛先アドレス (宛先エッジルーターアドレスを記入)				
オプション			パディング	

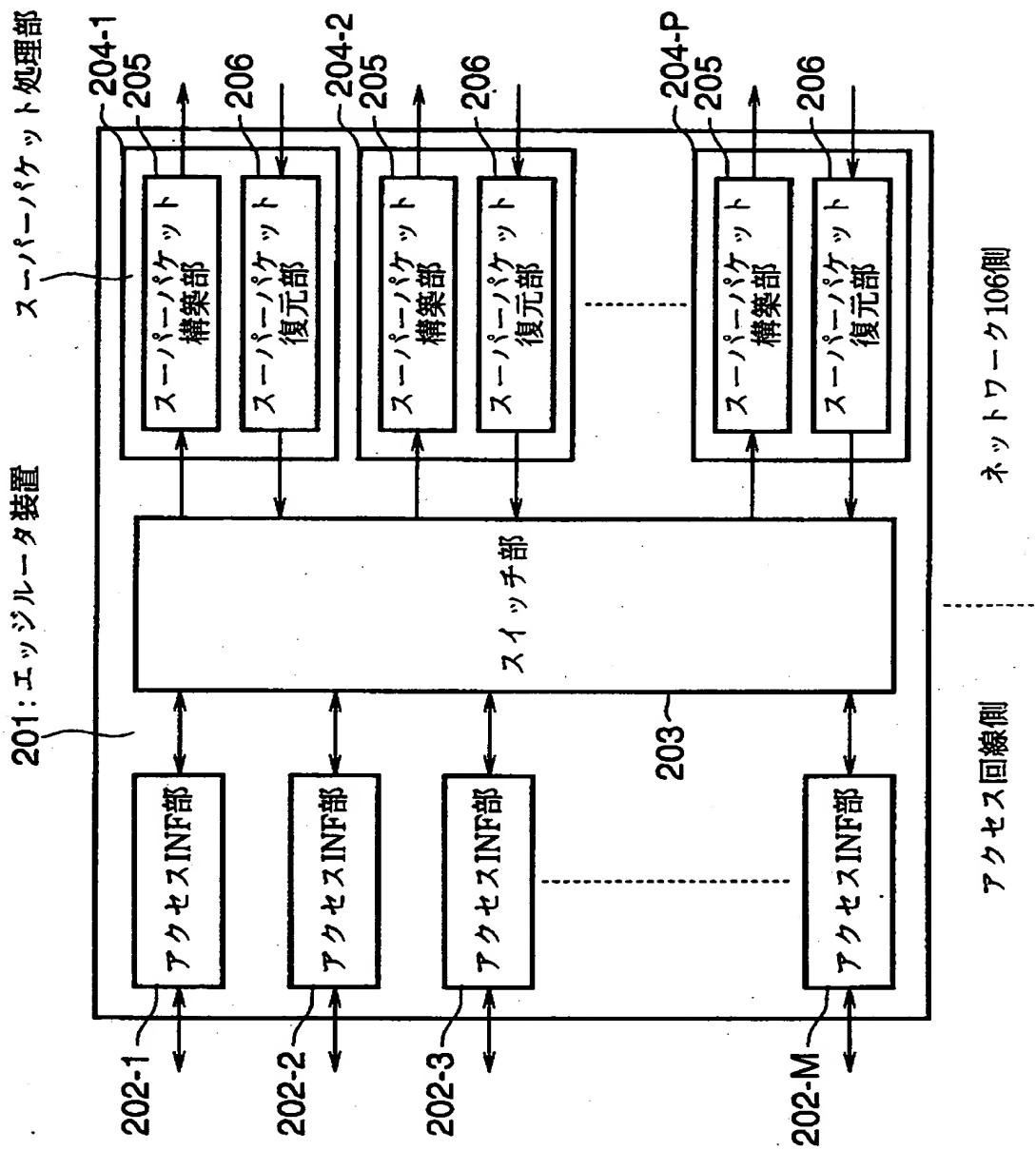
【図 5】



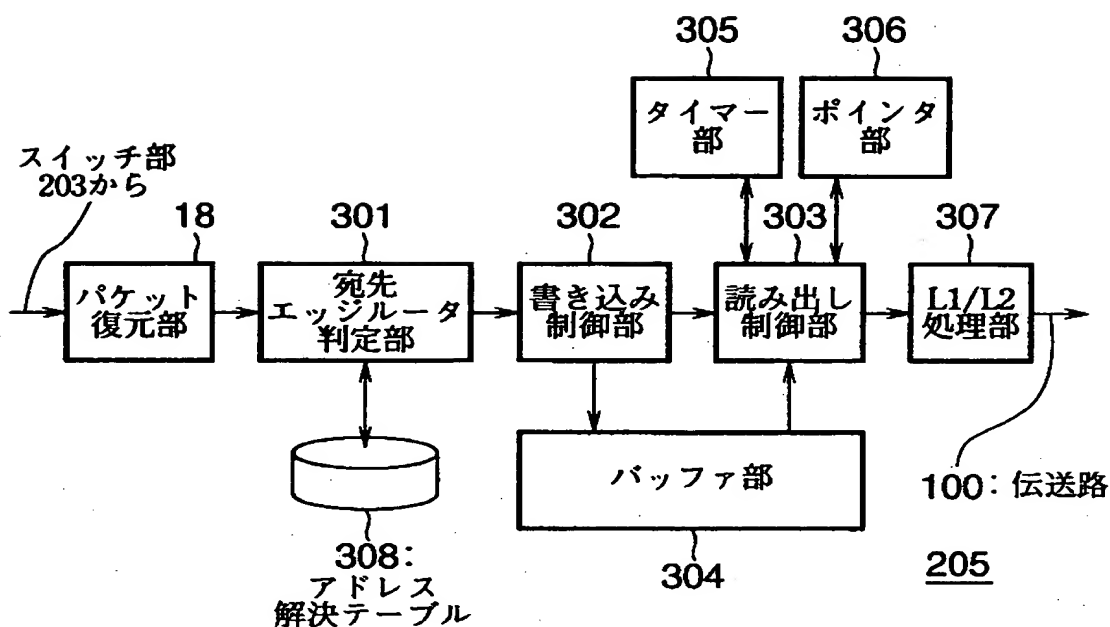
【図 6】



【図 7】



【図 8】



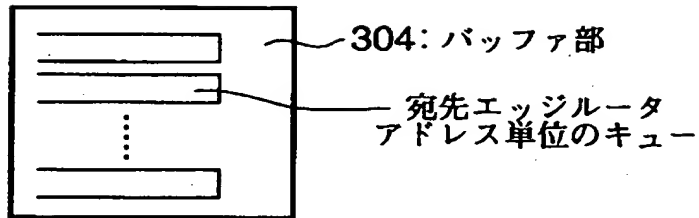
【図 9】

各Entry

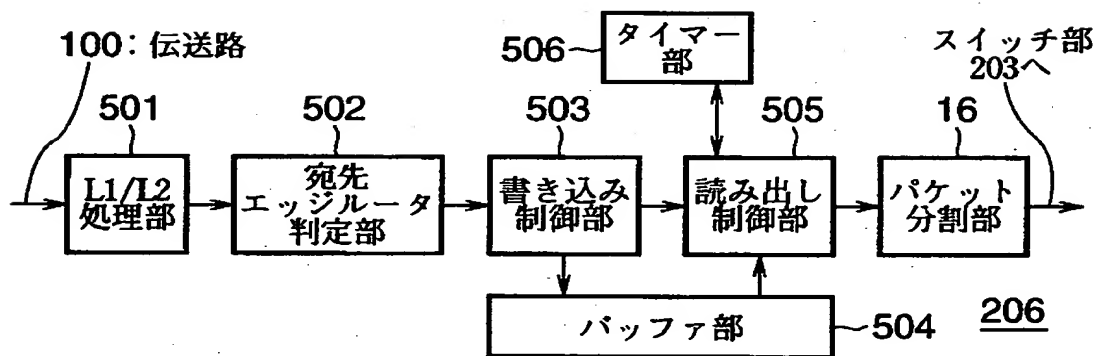
32bitの宛先IPアドレス

X1. Y1. K1. Z1	Egress エッジルータの 網内アドレス	Binary Tree内における 隣接Entryとの接続を 示すポインタ情報
X2. Y2. K2. Z2	Egress エッジルータの 網内アドレス	Binary Tree内における 隣接Entryとの接続を 示すポインタ情報
X3. Y3. K3. Z3	Egress エッジルータの 網内アドレス	Binary Tree内における 隣接Entryとの接続を 示すポインタ情報
X4. Y4. K4. Z4	Egress エッジルータの 網内アドレス	Binary Tree内における 隣接Entryとの接続を 示すポインタ情報

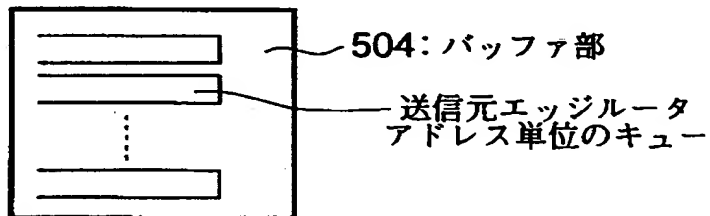
【図 1 0】



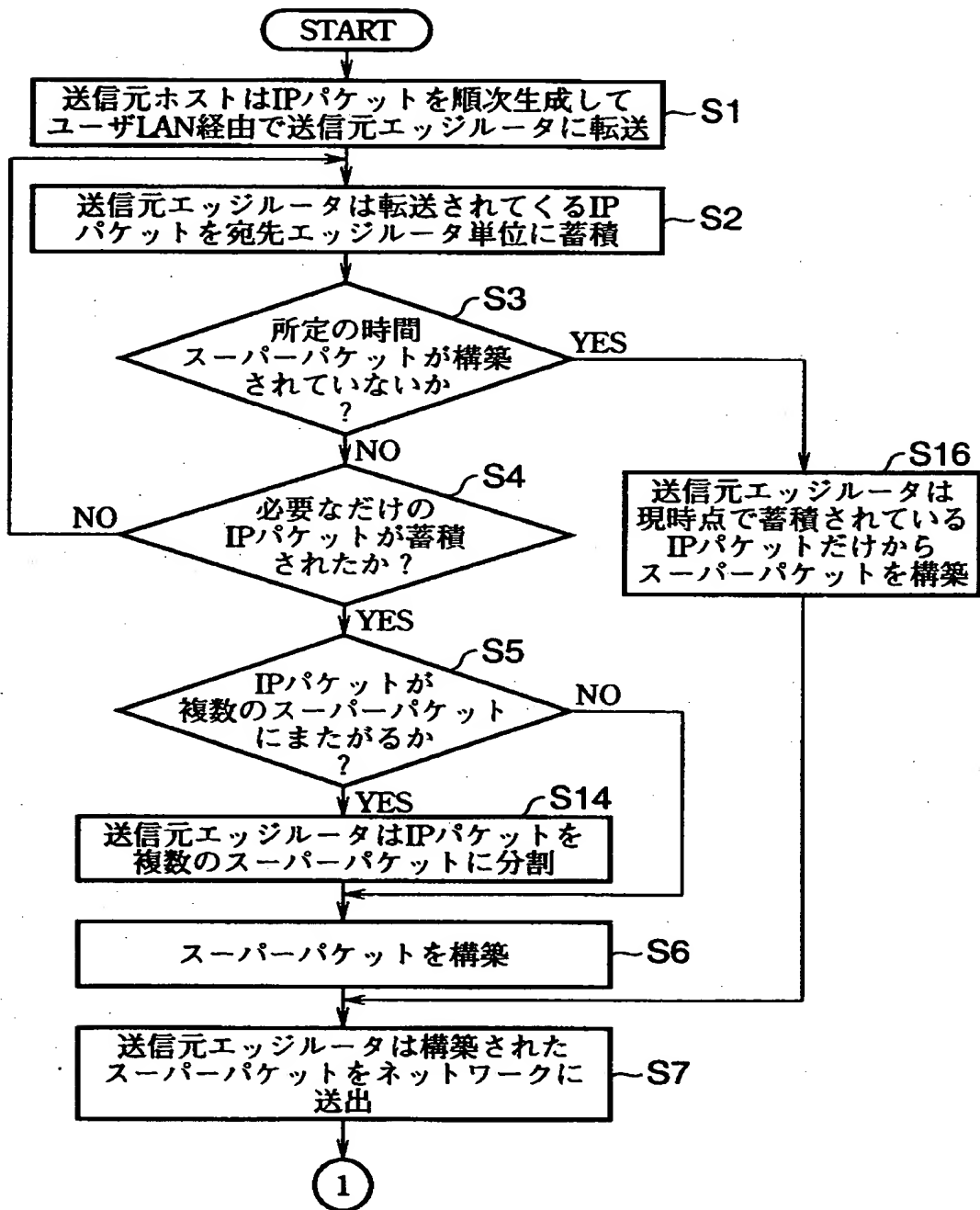
【図 1 1】



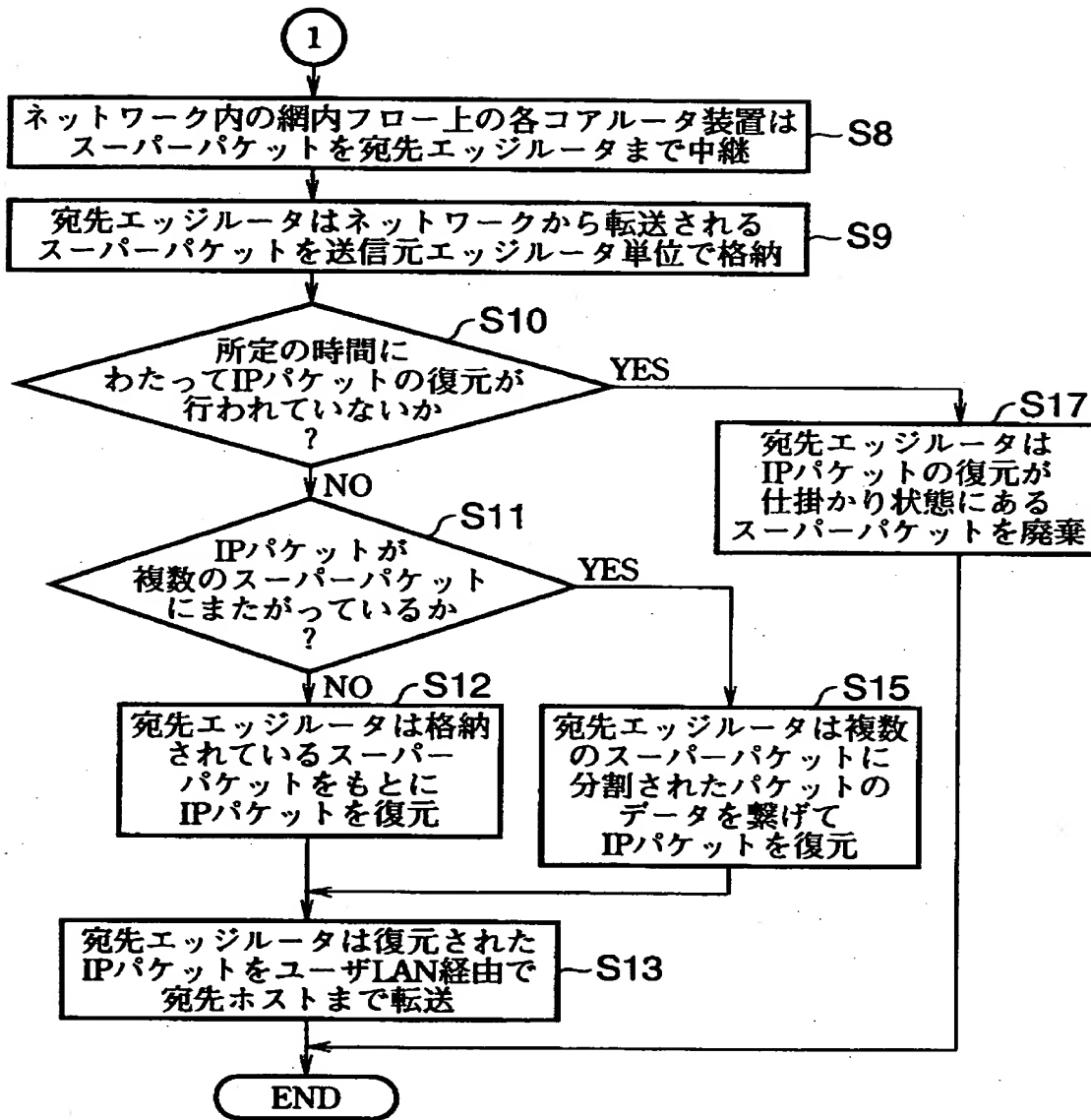
【図 1 2】



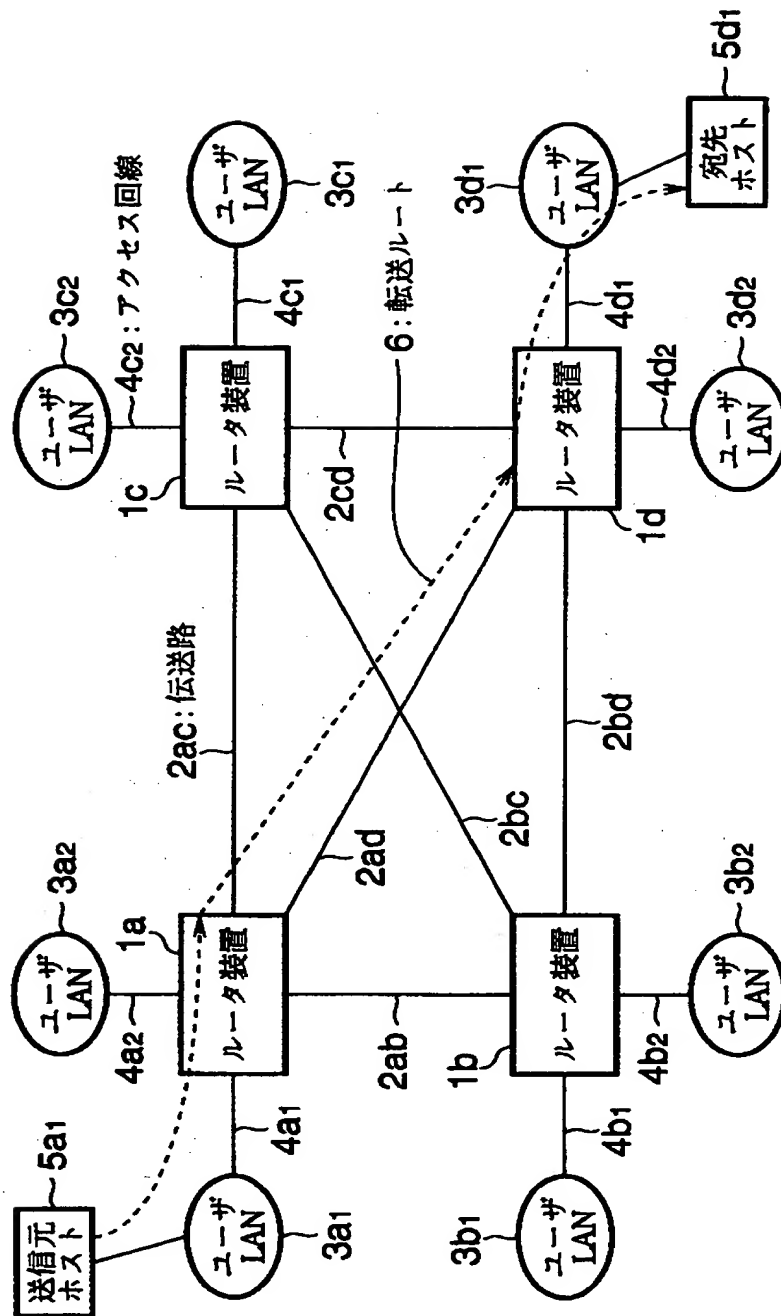
【図 13】



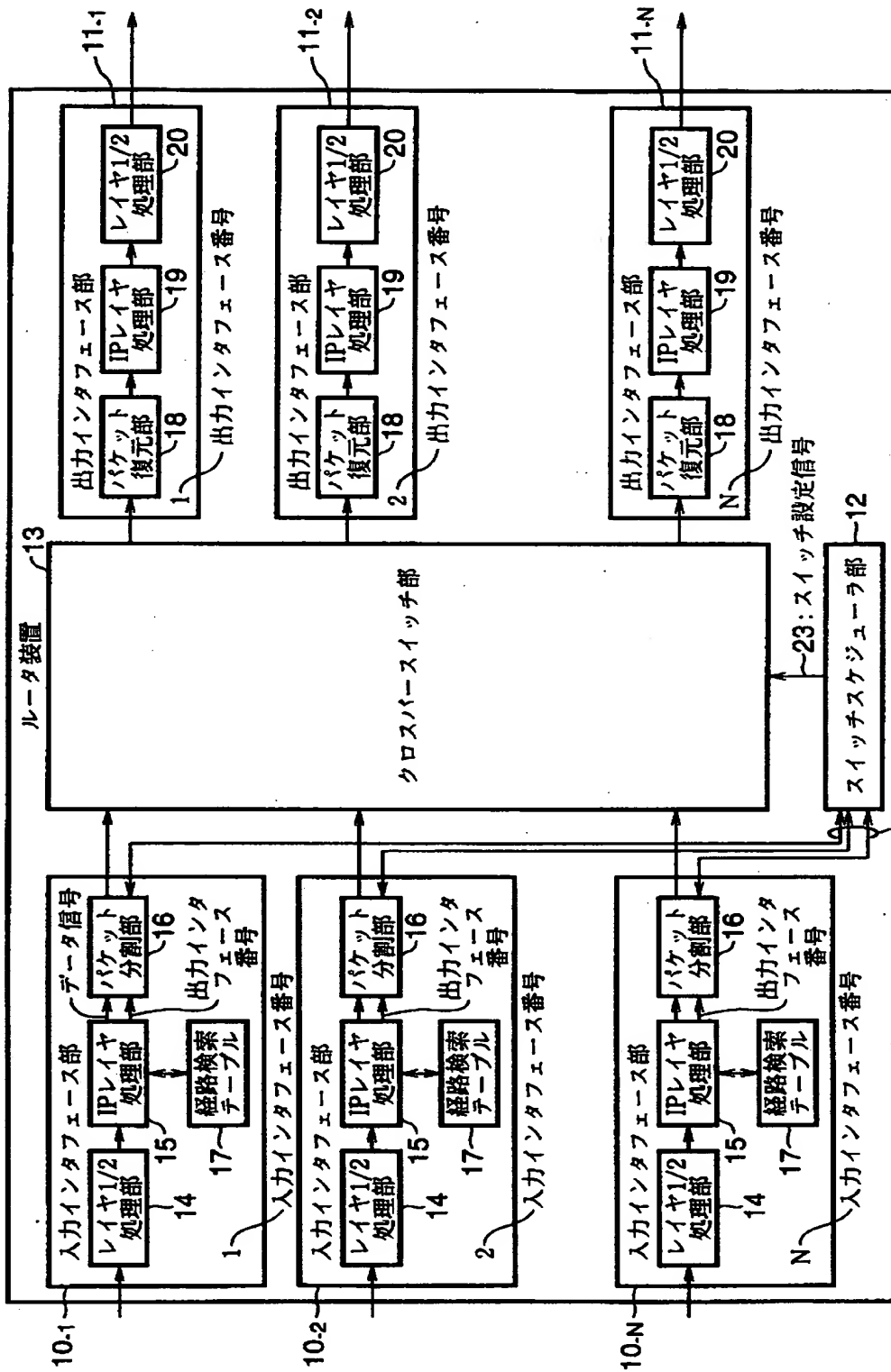
【図 14】



【図 15】

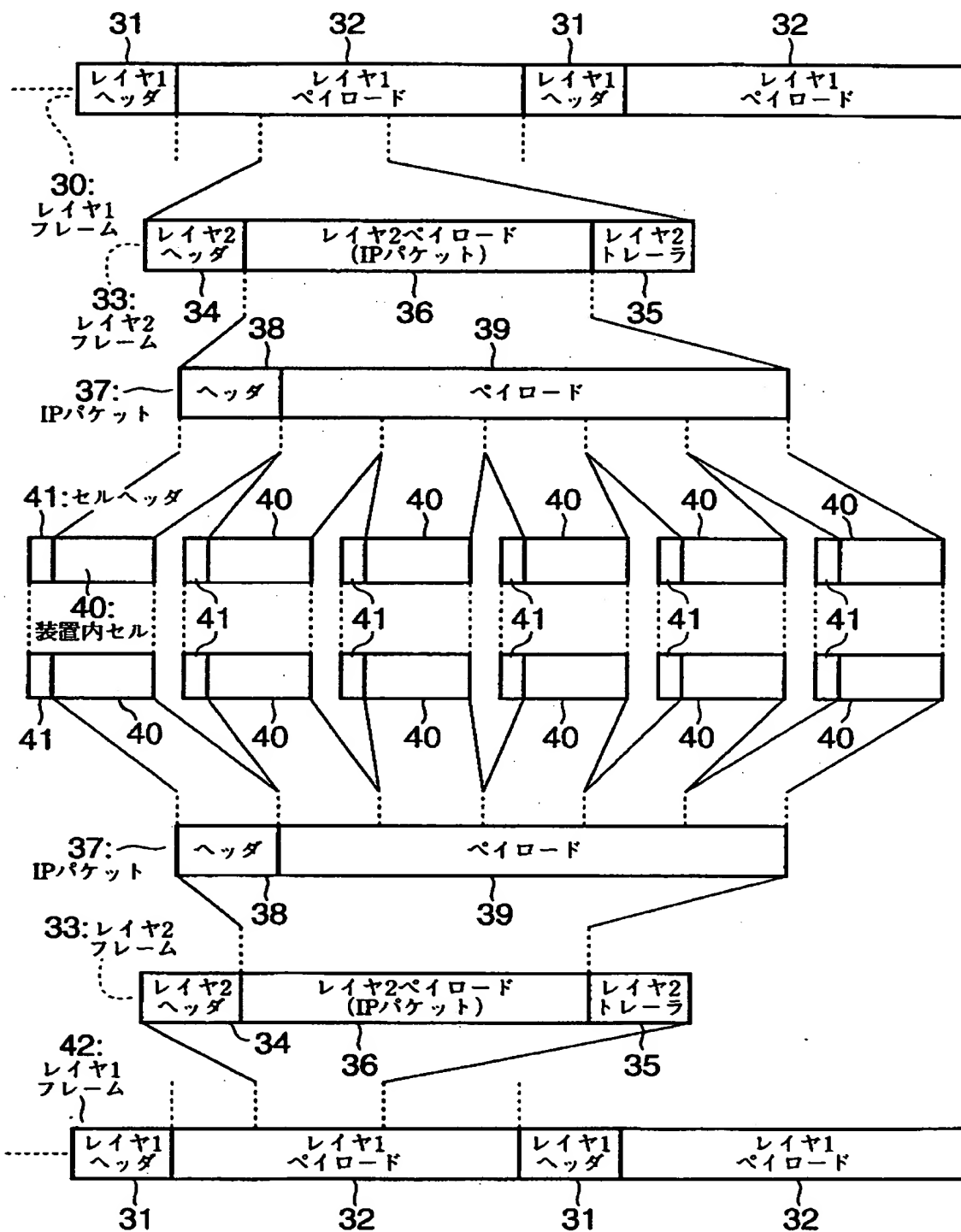


【図 1 6】

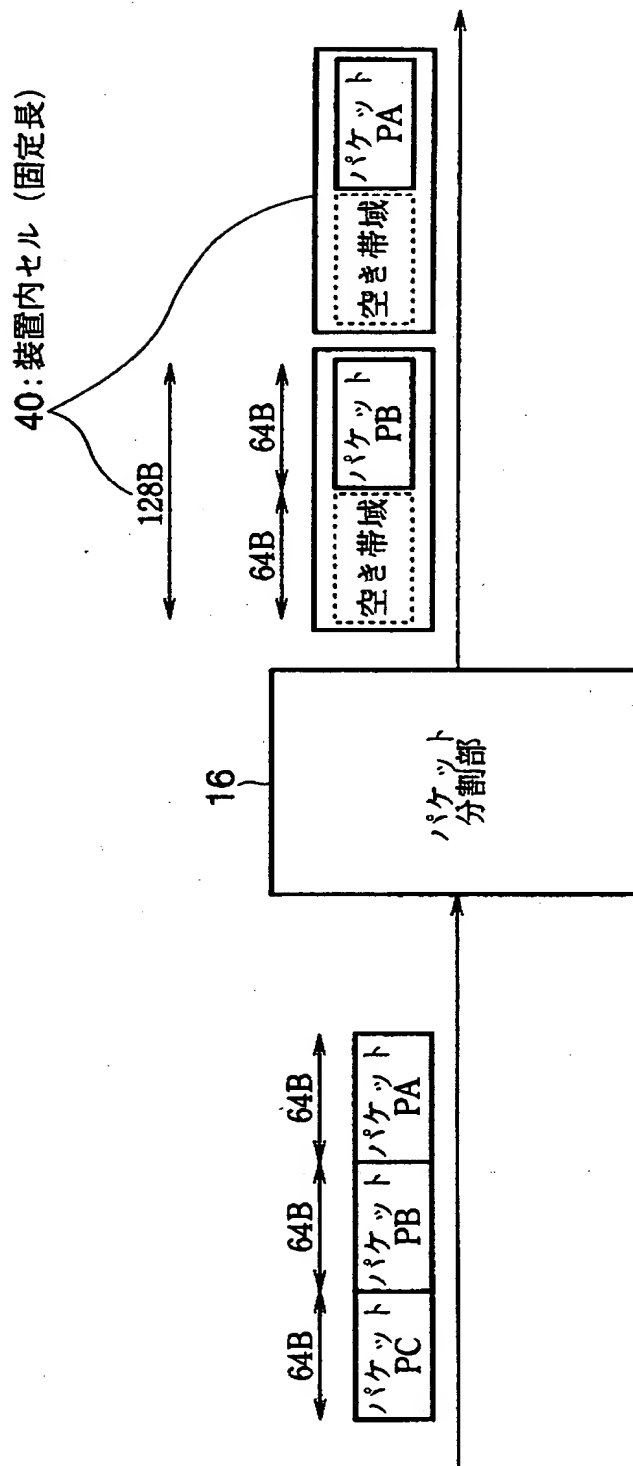


21: 転送リクエスト信号 (パケット分割部→スイッチスケジューラ部)
 22: 転送リプライ信号 (スイッチスケジューラ部→パケット分割部)

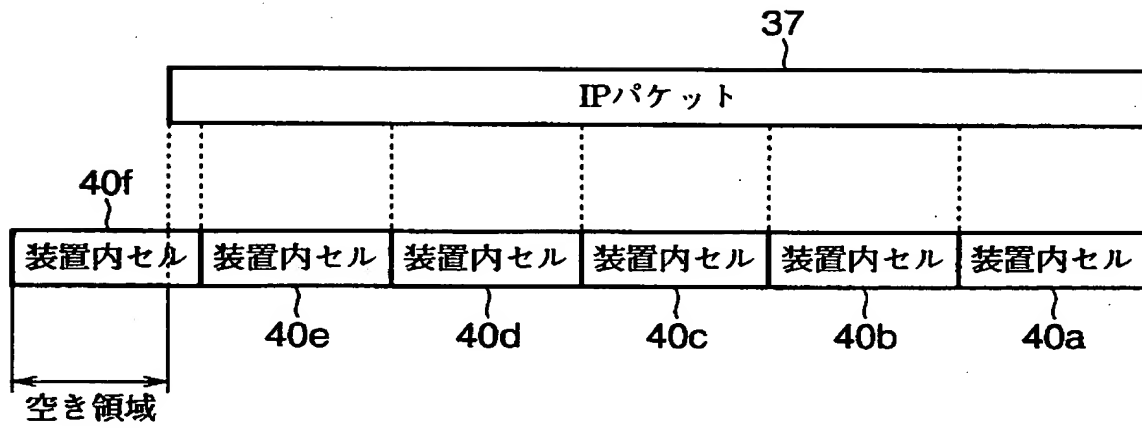
【図 17】



【図 1 8】



【図 1 9】



【書類名】 要約書

【要約】

【課題】 スイッチ容量を大幅に増やしてネットワーク上で転送可能なパケット数を増やせる大規模なパケット転送方法、装置及び通信システムを提供する。

【解決手段】 コアルータ装置はパケットを固定長Lの装置内セルにしてスイッチングする通常のルータである。送信元ホスト105cは可変長のIPパケット37をユーザLAN104c、アクセス回線103cからエッジルータ装置101cに送る。各エッジルータ装置はユーザLANとネットワーク106の境界に配置され、固定長Lの整数倍のスーパーパケット110とIPパケット37の間を変換してスイッチングする。スーパーパケット110は各ルータ装置に付与されたネットワーク106内固有の網内アドレスに従ってコアルータ装置102b、102dからエッジルータ装置101gに中継されてIPパケットに変換され、アクセス回線103g、ユーザLAN104gから宛先ホスト105gに転送される。

【選択図】 図1

認定 - 付加情報

特許出願の番号	平成 11 年 特許願 第 154308 号
受付番号	59900516017
書類名	特許願
担当官	木村 勝美 8848
作成日	平成 11 年 6 月 21 日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000004237
【住所又は居所】	東京都港区芝五丁目 7 番 1 号
【氏名又は名称】	日本電気株式会社

【代理人】

申請人	
【識別番号】	100108578
【住所又は居所】	東京都新宿区高田馬場 3 丁目 23 番 3 号 ORビ ル 志賀国際特許事務所
【氏名又は名称】	高橋 詔男

【代理人】

【識別番号】	100064908
【住所又は居所】	東京都新宿区高田馬場 3 丁目 23 番 3 号 ORビ ル 志賀国際特許事務所
【氏名又は名称】	志賀 正武

【選任した代理人】

【識別番号】	100101465
【住所又は居所】	東京都新宿区高田馬場 3 丁目 23 番 3 号 ORビ ル 志賀国際特許事務所
【氏名又は名称】	青山 正和

【選任した代理人】

【識別番号】	100108453
【住所又は居所】	東京都新宿区高田馬場 3 丁目 23 番 3 号 ORビ ル 志賀国際特許事務所
【氏名又は名称】	村山 靖彦

【選任した代理人】

【識別番号】	100100077
【住所又は居所】	東京都新宿区高田馬場 3 丁目 23 番 3 号 ORビ ル 志賀国際特許事務所

次頁有

認定・付加情報（続き）

【氏名又は名称】 大場 充

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社